

Omelettes without broken eggs?

The British government has reacted with hysteria to the indiscretions of a junior minister and to the wishes of its voters that the genus salmonella would go away.

ONCE upon a time, there was an engaging junior minister in the British government with a flair for dramatizing issues in public health and for irritating more senior and staid colleagues by frequently capturing the attention of the newspapers. These circumstances were always the more mystifying because the junior minister hardly ever strayed far beyond statements of the obvious — cigarette-smoking can damage health, a diet of unrestricted fat may contribute to the risk of coronary disease and so on. The minister was able so successfully to arouse opinion only by being so junior as to be free from the constraints that prevent wiser and more august colleague-statesmen from uttering even platitudes for fear of offending one or other of the vested interests whose opinions must be balanced in framing the compromises known as policy decisions and only by being so insignificantly lowly that people took no notice, continuing to smoke cigarettes and to exist on a diet of potatoes fried in deep fat. But then, one day, the junior minister foolishly said something that seemed to change people's eating habits, giving great offence to suppliers of the suddenly neglected food, and was forced to quit. The moral is that even ministers must learn that it is more important to be circumspect than to be listened to.

This, of course, is the sad tale of Mrs Edwina Currie, whose career as junior health minister ended abruptly last week after she had told a television interviewer that "most egg production" in Britain is affected by *Salmonella enteritidis*, the salmonella serovar most often infecting poultry stocks in Northern Europe. The consequences have been dramatic. Egg consumption has fallen by about a half from 30 million eggs a day, egg producers have been issuing writs against the departed minister threatening legal suits for the damage done to their businesses, the British government has advertised in the press its advice that eating eggs is safe provided that they are well (perhaps unpalatably) cooked, but that people should take extra care, especially about raw eggs, as in mayonnaise, and, finally, the government has intervened in what it calls an "abnormal market" by offering to buy eggs that would otherwise be unsold.

After the egg producers' pocket-books, and the romantic notion that a chicken's egg is a naturally sealed and aseptic package of nourishment (cholesterol included), public enlightenment is the most probable casualty of this frantic episode. For while Currie may have been technically correct, on one reading of her statement, and while nobody disputes that salmonella strains are a growing cause of food poisoning throughout Europe and North America, there are dangerous signs that Britain is about to give salmonellae the pride of place in the demonology of environmental hazards accorded, only a few weeks ago, to the more substantial business of climatic change consequent on the greenhouse effect. That would be a great misfortune.

Examinations

The reading of Currie's remark that rings true is that in which "most egg production" is understood to mean that some carriers of salmonella can be found in "most flocks of egg-producing chickens". Chicken farmers are required to carry out regular bacteriological examinations of their facilities for the agents of

zoonoses, and to report them to the Central Veterinary Service. By all accounts, the data confirm what an expert committee of the World Health Organisation reported a few weeks ago, in a document on *Salmonellosis control* (WHO Technical Report Series 774), that several of the 2,200 known serotypes of salmonella are increasingly troublesome pathogens of human beings which are carried, often benignly, by domestic animals. The trouble is that it is not possible to infer, from the isolation of salmonella from, say, faeces collected from a chicken-house, what proportion of its occupants may be infected. Even the fact that 60 per cent of chicken carcasses collected from British retail outlets carry salmonella bacteria does not imply a similar proportion of infected chickens in the flocks from which they came; killing chickens is too often a dirty business.

Evidence

At the same time, there is ample evidence of an increase of the frequency with which people are affected by food poisoning caused by chicken-specific strains. The Ministry of Health has logged 49 outbreaks (involving two or more people) affecting more than 1,000 people so far this year, while it seems that chicken-related salmonella infection has grown in a few years to account for more than half of nearly 20,000 bacterial isolates from individual food-poisoning cases. Given that food-poisoning, most of which tends to be unrecognized for what it is, is systematically under-reported, it is clear that Britain, which has had a long flirtation with salmonella poisoning, traditionally in pre-cooked meat pies, but more recently and conspicuously in aspic jelly supplied to an international airline, now has a novel version of a long-standing problem to contend with. If Currie had amended her statement to say just that, she might still have been a minister.

The British agriculture ministry appears now to acknowledge the force of what Currie might have said, and to be engaged on an exercise to limit the damage done to the egg industry without much thought for the long term. A code of practice for chicken breeders, on the stocks for the past six months and published at the beginning of the month, has been followed, in the past week, by one for egg-producers. No doubt there will soon be others. The recommendations so far rely on decontamination and bacterial examination, with the implied threat that veterinary officers may temporarily close facilities at which infection amounts to infestation. The guidelines (which are voluntary) do not commend vaccination (practised in West Germany). Both health and agriculture ministries seem to have been particularly disconcerted by the finding earlier this year of chickens with infected oviducts, suggesting (but not proving) congenital transmission from one generation to another, whence the requirement that egg-producers should systematically cull birds from their flocks and send their carcasses for bacterial examination.

In what seems certain to be a frenzy of activity ostensibly aimed at reducing the presence of salmonella in chickens and their products, two uncertainties will persist. First, there can be no way of telling when, or whether, Britain will again consume 30 million eggs a day. And, as things are, there will be no criteria for telling when, or whether, the measures now being taken to



reduce salmonella infection have been successful. The underlying difficulty is that salmonellae, while ubiquitous, are mostly at present cryptic. Telling whether there are bacteria in an egg, a carcass or in excreta requires that somebody should deliberately decide that some specimen or other should be sent for examination. (There is a danger that the present excitement and the requirements of the new codes of practice will quickly exhaust the available resources for bacteriological examination.) The cost of all this, in the months ahead, will be considerable. Would it not be better that there should be, instead, a programme of research designed to tackle some practical questions (How to diagnose infection more quickly? How to devise broad-spectrum vaccines?) while throwing light on more teasing conundrums?

The plain truth is that salmonellae, always with us, are here to stay. What people need in their defence against serious infection are ways of understanding the epidemiology and evolutionary biology of the bacteria. (Simple hygiene would also help.) It would plainly be a public service if there were a way of spotting in advance the emergence of bacterial attributes that are especially likely to kill people. In its response to Currie's indiscretion, the British government seems bent on mounting what looks like a campaign of salmonella eradication. Ironically, the erstwhile egg-eating population of Britain would probably be as much reassured by knowing about the uncertainties, and the prospects for their removal, as by the certainty, the way the wind is blowing, that British egg-producers will be testing everything in sight — until the demand for eggs picks up again.

What this implies is an urgent need for a more measured perspective than is likely to follow from the British government's alarm about the egg market. It would serve the public interest if some politician stood and up said that the salmonellae have been with us all for a long time, and are unlikely to be got rid of soon. Some information about the far from negligible chances of being infected except by eating eggs would also help. The news that people elsewhere are bothered by similar organisms in other domesticated animals (such as beef cattle) would paradoxically calm people's anxieties. Fears, substantial though they are, that developing societies are likely to be more afflicted would work in the other direction. □

Bush's budget problems

The new US president should worry about the bills that go unpaid as well as the federal deficit.

THE next president of the United States, Mr George Bush, may be at risk of prematurely exhausting the traditional indulgence that presidential newcomers enjoy in the minds of those who have elected them to office. Bush himself is not to blame. The circumstances of the succession (an incumbent vice-president succeeding a now-recumbent president) are responsible. The result is that Bush appears to be the man already in charge, although propriety requires that he should do nothing more substantial than announce (to general applause) who his colleagues will be. A month from now, when Bush's new administration will have joined battle with the Congress on the budget for the financial year beginning next October, the readiness to give him the benefit of the doubt may have evaporated. That will be a misfortune for Bush, but perhaps for the rest of us as well.

The problem of the US budget is now familiar: it does not balance. The reasons are simply understood. The federal government now spends more than \$ 10^{12} a year — one trillion dollars, or \$1,000 billion in American usage. This vast sum consists of three roughly equal parts — defence (\$300,000 million), statutory expenditure on what are chiefly social services (more than \$400,000 million) and the rest, ostensibly in the gift of the administration and the Congress, but including spending on schemes for providing the elderly with medical care

that may be abandoned only by exciting fierce political opposition. The federal deficit (this year likely to exceed \$140,000 million) is thus itself roughly a third of the federal government's apparently discretionary spending, which in itself explains why its modest reduction during the past few years has been so slow.

The argument that will begin in January has already been well rehearsed. How much, how quickly, can be quarried from the defence budget? What is the chance that doing nothing will cure the deficit because continuing prosperity brings higher incomes and disproportionately higher income taxes? What does Bush mean by a "flexible freeze"? Would a sales tax on gasoline count as a violation of promises that taxes will not go up? Even six months from now, these arid issues are unlikely to have been settled. The Democratic majorities in the Congress will have a vested interest in making the new administration swallow Bush's undertaking that taxes will not be increased, but nobody will wish to reduce spending. The potentially disastrous outcome is that issues that should be tackled, but which have been put off because of the budget deficit, will continue to wither.

Two obvious casualties of this conspiracy at neglect are the health of the US banking system (whose most lowly component, the savings and loans industry, needs between \$50,000 and \$100,000 million of new money) and the defence production reactors operated by the US Department of Energy, on which an estimated \$40,000 million must be spent in the next few years. With these clamant bills to pay, the US government is hardly in a mood to tackle problems that are more distant in the sense that their neglect will bring trouble only after the elected terms of present office-holders have elapsed. The most urgent of these is the appalling condition of public education in the United States, which has been amply catalogued during the past decade. It is mystifying that such an issue, which is widely acknowledged to be crucial to the industrial competitiveness of the United States (last year's good cause) and thus to the readiness of people elsewhere to continue financing the federal deficit, should be given so little attention.

The federal government, of course, will say that education is the responsibility of the fifty states, but that is a half-truth only. On at least two matters, the federal government's influence is decisive. First, by the provision of research grants through half a dozen agencies, the federal government helps to determine the quality of education at a small proportion of the institutes of higher education in the United States. In the battles that lie ahead, these budgets will be threatened with cuts when they should be being enlarged. Second, by its support of schemes for lending money to students for tuition payments, the federal government vitally influences access to higher education, but insufficiently. Much too little is being done to cast the net of professional training beyond those sections of the US population in which the benefits of professional training are already ingrained. While the traditional isolation of small-town rural America from wider intellectual currents may be eroded in coming decades, what benefits will that bring if the aspirations of the urban ghettos remain as impoverished as they are?

But higher education is not, in any case, conspicuously a disaster. Elementary and secondary education, especially in most major cities, is where the rot begins. The federal budget includes a few hundred million dollars for research and demonstration projects in these fields, but there is only a meagre chance that state and city governments will follow obvious precepts of good practice until the federal government is willing and able to persuade them in the right direction by spending matching funds. Nobody should be surprised that next month's budget (presented by the outgoing president and likely to be taken as a basis for negotiation by his successor) will not ask for funds for purposes such as these. The deficit is too big, and the need to reduce it too urgent. But that is a vivid illustration of how the deficit is excluding other good causes from the political agenda of the federal government. □

Privatization not the answer to NIH problems

■ High ratings for NIH research

■ Administrative, salary problems remain

Washington

THE controversial suggestion that the National Institutes of Health (NIH) research facility in Bethesda, Maryland, would be better off managed by a private enterprise is not justified, but some changes "are absolutely necessary if the program is to continue to be an important component of the nation's biomedical research effort". That is the conclusion of an Institute of Medicine (IOM) report* released this week that looked at a potential cure for some of the ills — real or imagined — that are plaguing NIH.

Although the White House Office of Management and Budget, where the idea first originated, quickly denied it was proposing selling off "a jewel in the crown", it nonetheless was anxious that the idea be given serious thought. The principal White House concern was that, given government wage limitations, NIH could not hope to retain top researchers. There had been reports in 1987 that AIDS researcher Robert Gallo was being wooed away from NIH by a private university.

But according to the IOM report, the intramural programme will not easily be privatized. NIH's Bethesda activities would not generate enough revenue from user fees or sales of services to support their research activities. Moreover, cutting the intramural programme off from the grant-giving extramural activities could destroy a unique element of the NIH research enterprise.

But the report is clear that there are serious problems confronting the intramural programme. Some 13,000 full-time employees work at the Bethesda research campus, as well as approximately 2,000 visiting researchers. The programme has an annual budget of around \$700 million, a little more than 10 per cent of the total for all NIH activities. Although salaries for junior researchers on average compare favourably with university and private industry, top researchers with either MDs or PhDs earn substantially less than their colleagues outside government. Top salaries rarely exceed \$100,000, often a starting point for industry or large university medical schools.

NIH also face other nagging problems. There is a widespread perception that the quality of junior scientists at NIH is slipping, perhaps because of competing offers from private industry. Government retirement programmes are incompatible with those outside government, making it hard

to recruit senior people. Research space is at a premium, and Congress has been reluctant to authorize new construction. International travel is controlled by the Department of Health and Human Services, the cabinet agency responsible for NIH, and NIH scientists generally believe their work is hampered by department restrictions.

Despite these problems, the IOM report judges that the intramural programme is still producing high quality research. A bibliometric analysis shows that in the period 1981 to 1984 more than one-fifth of the 10 most highly cited research papers were by NIH scientists. (NIH rank slightly below the Scripps Clinic and Research Foundations, and more substantially below Rockefeller University, in average number of citations, but such comparisons must be treated with caution.) There are also four Nobel laureates at NIH, as well as more than a dozen Lasker Prize winners and some 60 members of the National Academy of Scientists.

Having concluded that NIH is better served by remaining part of the government, the IOM report considers several options for improving the current situation. The report encourages the creation of a personnel demonstration project that would permit more flexible hiring and pay practices, and remove federal pay ceilings where justifiable. The report also encourages Congress to permit the creation of ten endowed chairs to attract or retain top scientists. Administrative problems can be ameliorated by giving the director of NIH greater autonomy for administrative decisions. Many of these decisions are at present subject to review from the Assistant Secretary for Health.

Other suggestions are to create a special director's discretionary fund of at least \$25 million to take advantage of unique research opportunities. External reviews conducted every four years of intramural programmes would help assure high research standards. Creating an NIH scholars programme for outstanding young investigators appointed on a competitive basis to non-tenured positions would help to attract a new cadre of scientists to NIH.

The IOM committee that wrote the report was chaired by Harold Shapiro, president of Princeton University.

Joseph Palca

* A healthy NIH intramural program: structural change or administrative remedies. National Academy Press, Washington DC, 1988.

Fang stays home

London

PROFESSOR Fang Li Zhi, the malcontent Chinese astronomer who was fired from his academic post three years ago, has now been prevented by the government of China from taking up an invitation to the United States.

Several colleagues in the West have written to *Nature* to protest at what appears to have been a sudden revocation of Fang's permission to travel abroad. As recently as October, he was in Australia for an international conference.

Fang's would-be hosts in the United States believe that the revocation of his permission to visit them stems from an interview with the Chinese press some weeks ago in which Fang, then in Hong Kong, offered the opinion that speech is less than free in mainland China.

Among those who have signed the letter of protest are John N. Bahcall (Princeton), George R. Blumenthal (Santa Cruz), George B. Field (Harvard), Riccardo Giacconi (Space Telescope Science Institute), Peter Goldreich (Caltech), Douglas N.C. Lin (Santa Cruz), Richard McCray (Colorado), Christopher McKee (Berkeley), Frank Shu (Berkeley), Harlan J. Smith (McDonald Observatory) and David N. Schramm (Chicago). □

Britain in Europe

London

LAST week saw the resolution of two separate conflicts between Britain and its European partners in collaborative research projects. Britain's science minister, Robert Jackson, said that Britain would remain a member of CERN, the European laboratory for particle physics near Geneva, having won substantial reductions in its future subscription. A change in the method of calculating subscriptions will mean that Britain's fee is no longer increased out of proportion by currency fluctuations. The new method will come into effect in 1990.

But Jackson said that Britain would be keeping up the pressure on CERN to improve management efficiency, reduce staff and to increase investment by non-member states, with a view to reducing the subscriptions of members. Britain's contribution will be reduced from £54 million this year to £47 million next year; Britain hopes to reduce that figure to £40 million in subsequent years.

Britain also finally agreed last week to the 5 per cent increase in the science budget of the European Space Agency, but only on condition that an independent review of the programme is carried out. This secures the immediate future of the Horizon 2000 programme which had been in doubt because of the British veto on increased support.

Christine McGourty

US and Soviet academicians meet in Washington

Washington

THE US National Academy of Sciences and the Academy of Sciences of the USSR last week announced plans for a joint committee to study the global ecology. The agreement coincided with a visit by a 16-member delegation from the Soviet Academy to the United States that was cut short by the earthquake in Armenia.

The impetus for the new committee, according to a report prepared earlier this month by a working group drawn from the two academies, was an awareness that the "problem of global ecological security that confronts all humankind has increased greatly in its seriousness". Increased use of natural resources, population growth and increased production of greenhouse gases all pose serious problems. "The global ecosystem is being degraded rapidly", according to the report, "so that its sustainability is being called into question." The new committee will meet once or twice a year to consider plans for studying the situation, and devising recommendations for steps to slow the deterioration. The committee will assist in bilateral activities, as well as with multinational efforts such as the International Geosphere/Biosphere Program.

The visit by the Soviet delegation began with a speech at the US academy's headquarters in Washington by Soviet academy president Guriy Marchuk. Marchuk spoke of how Mikhail Gorbachev's plans for restructuring Soviet society — *perestroika* — will affect the scientific community.

In practical terms, Marchuk pointed to mandatory retirement for academicians at age 75. Although they will retain most membership rights and perquisites, academicians over 75 will no longer hold administrative posts. There are also competitive elections for the leadership posts at scientific institutes.

Marchuk says that half of Soviet institutes are now headed by "young people". He went on to suggest that the best ideas usually come to scientists before they are 30 years old. Perhaps noticing the expressions on the faces of his audience, Marchuk quickly added, "Of course, there are some exceptions here".

Marchuk called the renewed emphasis on economics one of the biggest changes wrought by *perestroika*. At first, Marchuk said, scientists were not prepared to find the tools to contribute to the process of *perestroika*, but, he said, that is now changing. He also stressed that *perestroika* is directed toward humanitarian goals and humanity. "Everything should be done for the benefit of man", Marchuk said.

Joseph Palca

New clinical trial programme for AIDS in the United States

Washington

IN an attempt to widen access to clinical trials of experimental AIDS treatments and bolster confidence in the clinical trial process, the US National Institute of Allergy and Infectious Diseases (NIAID) late last month announced the creation of a new community-based clinical trial programme. Under the new programme, community physicians and clinics will systematically monitor their patients who are receiving experimental therapies and report the results to NIAID. In this way, NIAID hopes to offer the benefits of new drugs to more minorities and intravenous drug abusers, who are often under-represented in clinical trials, and encourage greater compliance with clinical trial rules.

Lack of adherence to the set protocol has been a problem in many clinical trials of AIDS drugs, especially when protocols called for control groups to receive only a placebo. People with AIDS often mistrust the motives of AIDS researchers, but join clinical trials because it is the best way to obtain promising drugs.

Some clinical trial participants, desperate for a cure, have supplemented the drug they were receiving with unproven drugs available on the black market, such as dextran sulphate and AL-721 (see *Nature* 335, 485; 6 October 1988), confounding the trial results.

Community-based organizations, such as the Community Research Initiative led by Tom Hannan in New York City, have had a greater level of compliance in clinical trials. But most of the community groups consist of middle-class homosexual men, who are better organized than the other groups most affected by the AIDS epidemic. Results obtained from trials involving primarily homosexual males may not apply to the population at large.

The NIAID programme will attempt to make use of existing community groups, while encouraging others to form. One resource that may be fruitfully exploited is drug-abuse treatment centres and public health clinics, where people with AIDS often enter the health-care system. Lawrence Deyton, the NIAID researcher who will be leading the project, says these centres will need guidance on how to study AIDS treatments.

To this end, part of the \$6 million devoted to the NIAID programme in 1989 will be used to assist community physicians to work out scientifically sound clinical research protocols.

Deyton says community groups and physicians will be able to apply for funds to study whatever drug they wish, but NIAID will guide the programme so that

valid data are collected. In some cases, Deyton says, the community-based trials could even substitute for some of the studies required by the US Food and Drug Administration (FDA) for approval of a drug. NIAID will supply the drug to be studied to the community group or physician.

Most of the large community groups formed in response to the AIDS epidemic are expected to support the programme. Martin Delaney of the San Francisco AIDS education group Project Inform says he intends to submit an application to participate. But he says the test of the programme's efficacy will be whether the money really goes to local community groups, and not to the university clinical research centres who are conducting most of the current clinical trials.

Deyton says he will require all applicants to the programme to provide letters of endorsement from the community they serve to prove that they are truly representative of the community's interests.

The panel which will review the community applications will also be composed of "true peers", Deyton says, including people with AIDS as well as scientists and pharmaceutical company representatives.

Perhaps the largest obstacle facing the programme will be the paucity of basic health care in the inner-city slums that are often hardest hit by the AIDS epidemic. Anthony Fauci, director of NIAID, says this problem is "even more complex than I would have imagined", because it points up basic faults in the US health-care system. Fauci says he will try to focus discussion within the government on how the US Public Health Service can better cope with the problems caused by AIDS.

Carol Ezzell

■ The US Food and Drug Administration last week approved a test kit for the presence of antibodies to HIV, the virus causing AIDS, that gives results in only five minutes. The test consists of envelope proteins from HIV bound to latex particles, which visibly agglutinate when antibody is present.

The test is intended primarily for use in physicians' offices, and will not replace the immunoassay tests now used to screen blood donated to blood banks because it has a higher rate of false positive results.

The HIV proteins used in the kit are produced by genetic engineering techniques by Cambridge BioScience, a Massachusetts biotechnology company. Cambridge BioScience licensed the patent for the gp120 envelope protein of HIV from Harvard University earlier this year (see *Nature* 331, 649; 1988). The kits will be sold by Baxter Health Care Corporation. □

Stock market scandal shakes telecommunications executives

Tokyo

THE chairman of Nippon Telegraph and Telephone (NTT), Japan's domestic telecommunications giant, last week became the latest victim of a stock market scandal that has shaken the Japanese government and industry.

The company at the centre of the stock scandal is Recruit Co., prominent in Japan's booming information industry. Between 1984 and 1986, Recruit sold millions of shares in Recruit Cosmos, a real-estate subsidiary, to prominent leaders in government, industry, universities and the media, before the shares became available to the general public. Many of

the buyers made large profits in October 1986, when the shares went on public sale at a premium.

Senior NTT officials feature prominently among recipients of the shares. It is widely believed that Recruit sold them the shares to win help in entering the lucrative telecommunications market following the privatization of NTT in 1985. Recruit leases high-speed digital communication lines from NTT and re-sells them to companies for video conferences and meetings at rates far below those of NTT.

Last week, Dr Hisashi Shinto, chairman of NTT, resigned after it was revealed that profits from the sale of Recruit Cosmos

shares had been deposited in his account. Shinto's resignation followed closely that of Kiichi Miyazawa, finance minister and deputy prime minister, who had to step down because of contradictory statements he made over his role in the Recruit affair.

Details of the profiteering began to leak out in July, shortly after a vice-mayor was fired for his role in the share deal. The scandal escalated when an official of Recruit was videotaped offering an envelope stuffed with cash to an opposition member of the Diet who was looking into the affair.

The videotape was subsequently broadcast on national television and the Recruit official has been indicted on charges of attempted bribery.

Former Prime Minister Yasuhiro Nakasone is one of the key figures in the Recruit scandal. Two of his secretaries and a former accountant of his supporters' organization bought a total of 29,000 Recruit Cosmos shares. And several cabinet members of Nakasone's former administration, including the present Prime Minister, Noboru Takeshita, and Shintaro Abe, secretary general of the ruling Liberal Democratic Party, have been alleged by the Japanese media to have links with the share dealing. But Nakasone is maintaining a strict policy of "no comment".

The resignations of Finance Minister Miyazawa and NTT chairman Shinto may have profound consequences for Japan. Miyazawa and Abe were the two strongest candidates to replace Takeshita when his two-year term as prime minister ends next year. But Miyazawa's hopes have now been dashed and his faction within the ruling party has been severely weakened. Shinto led the privatization of NTT with Nakasone's backing, and his resignation may result in greater pressure from the Ministry of Post and Telecommunications to break up the telecommunications monopoly, which still controls all but a few per cent of the domestic market.

But in the long run, the scandal may more seriously undermine the reputation of Japan's government and industry with partners overseas just at the time when both are seeking greater influence.

David Swinbanks

Medicines committees disclose links with drug companies

London

FOR the first time, the British government has revealed details of the personal financial links between members of the bodies that oversee the safety and licensing of drugs and the pharmaceutical industry. The links are extensive. But the government maintains that commercial interests do not influence the advice it receives from these bodies. It is making the details public simply "to avoid public concern that the interests might influence their decisions". But publication seems likely to have the opposite effect.

Under the ethical code to which members must adhere, consultancies, fee-paid work or shareholdings must be declared.

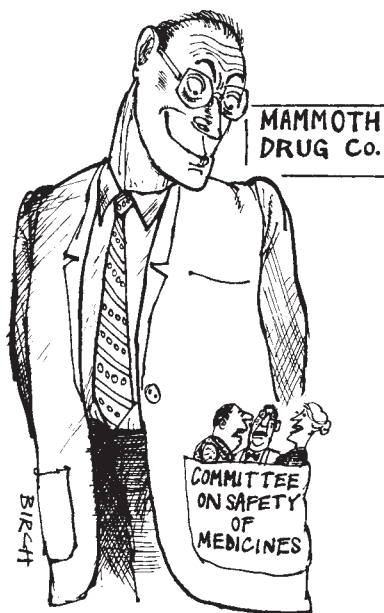
Of the 21 members of the Committee on Safety of Medicines, which advises the government on licensing drugs and monitoring those on the market, 13 have declared interests. And of the 24 members of the Medicines Commission, 16 have declared interests.

The ethical code requires that if a member has any interest in a product under consideration, that member cannot take part in the proceedings except, at the chairman's discretion, to answer questions by other members.

The Minister for Health, David Mellor, says that having links with industry helps members to keep in touch with important aspects of drug development. And in the ethical code, too, it is stated that it is desirable that members have an understanding of the work of industry and practical experience of scientific problems of product development.

Next year will bring more *glasnost* within the drugs watchdogs: indirect links with industry, such as the sponsoring of

research in a member's department, will also be made public. But there the openness will end; a spokesman for the Department of Health says it will not give in to demands that the amounts of money involved be disclosed or that members should be made to sever financial links



with industry while advising the government. Only with these concessions can advice to the government be impartial, says Dr Joe Collier, a pharmacologist at St George's Hospital Medical School in London, who has campaigned for the register of interests to be made public. He says that by requiring that the chairmen of the commission and the committees sever all links with industry, the government is acknowledging that such links do influence decision-making. Christine McGourty

Living coelacanths

London

THE living coelacanth *Latimeria chalumnae* was discovered 50 years ago this week, on 23 December 1938 (see page 727 of this issue). A special exhibition to mark the occasion opens on 22 December at the British Museum (Natural History) in London. Models and memorabilia of the discovery will be on display, and a video film of living coelacanths in their natural habitat will be shown.

Henry Gee

Emergency aid voted for West German universities

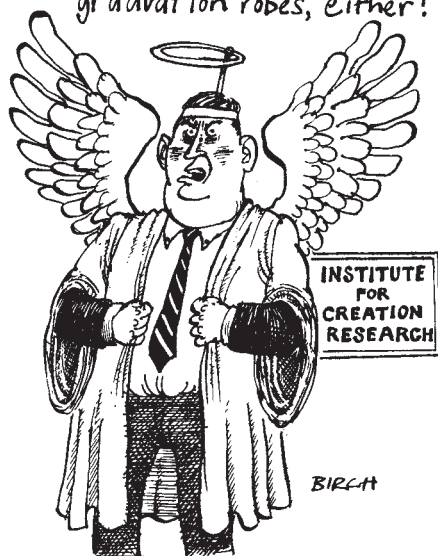
Munich

WEST German politicians last week approved an emergency aid package to relieve university overcrowding in certain highly popular fields just as a wave of student demonstrations reached a new high-water mark.

Chancellor Helmut Kohl and the minister-presidents of the 11 *Länder* (states) agreed on 15 December that federal and *Länder* governments would each contribute half of Dm2,100 million in university support spread over seven years beginning in 1989. The package had been suggested by West German Education Minister Jürgen Möllemann.

The programme details were left unclear, but much of the money is expected to be spent on creating a larger capacity in

They won't allow our graduation robes, either!



popular fields such as business studies and information sciences.

The emergency aid follows weeks of demonstrations in which tens of thousands of students have boycotted classes and taken to the streets. The protests, which began at a few big-city universities such as Frankfurt, Cologne and the Free University of Berlin, have spread to as many as 40 universities. They were scheduled to continue at some places until at least 20 December.

As each *Länder* has autonomy in deciding university policy, the student protests have focused on different issues around the country. But almost everywhere there are complaints about housing shortages, overcrowded lecture halls, unbearable conditions and a marked decline in the quality of education.

The chronic overcrowding that has become part and parcel of university

education in the 1980s grew noticeably worse this semester.

In North Rhine-Westphalia, student anger was directed at the *Land* government for shifting positions away from the liberal arts and into natural sciences and engineering. In other *Länder*, there were calls for more student involvement in university affairs and new chairs in women's studies.

What made the protests unusual was the encouragement and support of university faculty and staff members. "Bad conditions for instruction of the students mean bad working conditions for the staff", said the employees' organization at the Free University of Berlin.

The demonstrations have differed dramatically from the famous 1968 student movement, when protesters sought to

change society and used universities as a platform to air their views. In 1988, students are concerned above all with the quality of their own education.

The West German Rectors' Conference (Westdeutsche Rektorenkonferenz or WRK), which represents the interests of universities, welcomed the programme as a step in the right direction, but warned that much more must be done before the real problems at universities can be solved. Investment in higher education has stagnated for a decade, said WRK president Hinrich Seidel, while the number of students has nearly doubled, to roughly 1.5 million this semester. The new money corresponds to just 2 per cent of the cost of running the universities, he said.

Seidel wants a "substantial portion" of the money to go to fields besides business and information sciences. He also expressed the fear that the new money might aggravate the economic differences between northern and southern *Länder*.

Steven Dickman

No graduate degrees for creationists

Berkeley

CREATIONISTS in California are complaining of religious persecution, after the state Department of Education denied a small creationist school the right to award post-graduate degrees in science.

The Institute for Creation Research in San Diego is best known for its faculty member, Duane Gish, a frequent participant in public 'creation versus evolution' debates. Only 12 students have graduated from the institute's graduate school since it was founded in 1981, but their masters' degrees in biology, geology and astrophysics make them eligible for certification as secondary school science teachers.

A committee of scientists and educators assigned to evaluate the programme as part of the state re-approval process found that the students were not learning the

fundamentals of science because of the institute's strong focus on the biblical interpretation of creation. William Rukeyser, spokesman for the state Department of Education, says the department would be derelict in its duty if it allowed the institute to continue to award science degrees.

State school superintendent Bill Honig has clashed with creationists before over their political influence on the treatment of evolution in biology textbooks for secondary schools. The creationists accuse him of having a personal grudge against them.

"As a private Christian institution, we ought to be able to teach what we want", says Henry Morris, president of the institute.

Honig says the school may issue degrees in creationism or religion, but not science. "It's a consumer issue", he says. "If a person is going to get a degree in anything, the institution should be approved for quality in that area." Morris argues that the institute's science courses are very similar to those taught elsewhere, "except that the framework, which comes up every once in a while, is creationist instead of evolutionist". But Honig argues that biology, geology and astrophysics are based on fundamental assumptions about evolution and the age of the Earth and Universe that are incompatible with creationism.

The evaluation committee first voted 3 to 2 to approve the institute's programmes, despite the lack of scientific substance. But Honig convinced them that they should change their minds. "You don't have academic freedom to tell the public you're teaching something when you're not teaching it at all", he said. Marcia Barinaga

Experts to Armenia

Washington

A TEAM of earthquake experts from US universities, private industry and government agencies is scheduled to arrive in Armenia this week to help Soviet colleagues respond to the recent earthquake there. The team will offer advice on reconstruction efforts, and study how buildings were affected during the earthquake. An analysis of how well search and rescue efforts were carried out will be used as a basis for future preparedness efforts.

The team leaders are seismologist John Filson of the US Geological Survey and Mikhran Agabian, chairman of the civil engineering department at the University of Southern California. Joseph Palca

Commons committee criticizes information technology spending

London

THE British government's failure to tackle effectively the balance-of-trade deficit in information technology (IT) was criticized last week in a scathing report by the House of Commons Trade and Industry Committee. It accuses the government of complacency about the negative trade balance, which stood at £2,226 million in 1987. The government does not attach high priority to IT as a frontier technology needing sponsorship, it says, and this attitude does not promote use of IT in the economy. More funding for IT research is recommended as well as measures to tackle the country's "deplorable" and "worsening" skills shortage in this field.

Applications for places on university and polytechnic computer science courses fell by 38 per cent last year and the Committee of Vice-Chancellors and Principals (CVCP) expects a shortfall of a quarter in the number of students for computer science courses in 1989-90; the number planned for was about 7,000. The shortage of skilled people at present is estimated to be about one-tenth of the total number of IT professionals in users' and suppliers' companies. The committee received evidence of 30,000 unfilled vacancies for people with IT skills. This situation is "deplorable", and will take time to redress as "serious inadequacies" in the schools and in higher education are unlikely to be resolved swiftly. The country faces a crisis in both quantity and quality of people with

But the best solution to the skills shortage is increased in-service training, says the committee. It calls on the government to prove with statistics that companies are now spending more money on training, as it says they are. And it recommends that expenditure on training should be disclosed in company accounts.

The country's comparatively poor IT performance is also linked to expenditure on research and development. The committee was told by the CVCP that the highest proportion of unfunded alphabetized proposals supported by the Science and Engineering Research Council (SERC) is in computer science. Worthwhile research is going unsupported and Britain is not maintaining the proportion of its gross domestic product spent on research and development, says the committee. It finds no clear rationale for how IT research priorities are chosen and how the amount of money spent on each is determined. And it recommends that the Department of Trade and Industry (DTI) should consider increasing by £10 or £20 million its annual expenditure on alphabetized computer science research projects and the accepted projects of the national programme.

In conjunction with Britain's contribution to the European Commission's Esprit 2 programme, the national programme is now the main thrust of the country's IT research, replacing the Alvey programme, and is funded by the DTI and SERC. Alvey was designed to ensure that the country's IT industry could compete with Japan and the United States in the world market. But the committee doubts whether this has been achieved.

To encourage industry to spend more, the committee suggests that the government should identify the level of research to be carried out in the IT industry so that companies could aspire to this standard. But it suggests that tax incentives be considered for research and development carried out in small and new companies.

And it recommends that value-added tax should be waived on donations of equipment and software given by companies to universities. The tax rules should also be altered so as not to discriminate against donations which are for advertising and philanthropic purposes.

The implementation of the single European market in 1992 could herald more problems for Britain's IT industry. There is a real danger, says the committee, that the opening up of the European market will create opportunities for Japanese and US companies that will not be balanced by equal access for European companies to those markets.

Christine McGourty

Coordinated US science budget

Washington

MORE attention needs to be paid to developing a "cross-agency perspective" to unify and direct US government spending on research, according to a report* released this week by a joint committee from the US National Academy of Sciences, the National Academy of Engineering and the Institute of Medicine.

The committee concluded that, although the budget process for supporting research within an individual agency is working adequately, more effort is needed to ensure that the total US expenditure in science and technology matches the country's needs and priorities.

The committee was asked by Congress in the summer, after the annual budget process had been completed, to examine how well the current budget process works. Congress was concerned that co-ordination of the \$60,000 million dedicated to science and technology in the federal budget — spread over more than 30 agencies and institutions — was becoming too cumbersome and disjointed.

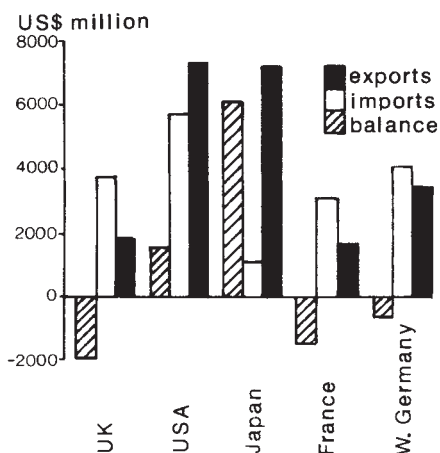
The joint committee recommends that the president, with the advice of his science adviser, should provide federal agencies with an outline of a national science agenda early in the budget cycle. Agency heads could then ask for money for inter-agency programmes without having to justify them exclusively within their own terms of reference. Supplied with a clearer picture of the funds needed by each agency to accomplish national goals, Congress would then be better equipped to judge the president's budget request.

The committee suggests that the science budget be broken down into four categories: research necessary to accomplish individual agency missions; inter-agency programmes which support the science base, such as education and basic research; projects important to national objectives, such as enhancing economic competitiveness; and major long-range plans, such as the space station or sequencing the human genome. Because the last three categories cut across agency boundaries, they would be evaluated by Congress against the president's overall plan.

The committee also asks that President-Elect George Bush should appoint a group of experts to determine how much of the defence budget can appropriately be included in estimates of US research efforts. At present, spending on defence research accounts for two thirds of the total federal spending for science

Carol Ezzell

*Federal Science and Technology Budget Priorities: New Perspectives and Procedures, 1988, is available from the National Academy Press, Washington, DC.



Trade balance in electronic data processing equipment in 1986. (Source: DTI.)

IT skills, says the committee.

One way to tackle this would be to increase the number of undergraduate and postgraduate courses in IT, and the committee recommends that a working party be set up urgently with the Department of Education and Science, the higher education funding councils and the business sector in order to do this.

Court ruling pleases UCSF

Berkeley

THE California Supreme Court relieved research universities of what they regarded as a potentially insurmountable burden in their continuing struggle with local residents seeking to block construction of new research facilities. The court ruled that the University of California at San Francisco (UCSF) did not have to prove that its new Laurel Heights building could not possibly represent a toxic hazard. Had the court decided otherwise, university officials throughout the state feared that they would be faced with an impossible burden of proof in similar cases.

UCSF bought its new research building in the Laurel Heights area of San Francisco in 1985, to relieve crowding on its main campus. Conversion of the space to research laboratories was halted in 1987 by a lawsuit filed by residential neighbours (see *Nature* 328, 656; 1987) who suspected that the new facility was potentially hazardous. UCSF had conducted environmental sampling studies on its main campus, which showed no increase in levels of toxic or radioactive chemicals in the vicinity of research buildings.

Those studies were used as evidence that there would be little risk at Laurel Heights. But local residents were satisfied, arguing that the absence of evidence that the emissions were harmful was not the same as proof that they are safe, and a local court agreed to stop the university's expansion plans.

The university appealed the case to the state supreme court, but in the meantime nearly 80,000 square feet of potential laboratory space remains empty, and the research groups already in the building are banned from using radioisotopes.

UCSF and other universities facing similar community opposition fear that the university would be given the impossible task of proving that all emissions were completely safe. "The scientific and biomedical world have been watching this case because of its implications for the future of research universities everywhere", said UCSF chancellor Julius Krevans.

Their fears were allayed with the court's acceptance of UCSF's risk assessment. But the university's troubles are not over. The court found that an environmental impact report that the university was required to produce did not adequately consider alternatives to the Laurel Heights building, nor did it adequately account for possible future uses of the building. A new report must be prepared and approved before the building conversion can continue.

Marcia Barinaga

Shortages of staff and money delay climate change predictions

London

BRITAIN'S research into the greenhouse effect could be rapidly speeded up with an increase of resources. At the rate research is now proceeding, reliable predictions of regional climate change could be 20 or 30 years away, but the same results could be available within ten years if more research funds were available, according to David Carson, director of atmospheric sciences, at the Natural Environment Research Council (NERC). The country's research effort into the greenhouse effect was criticized last week by the science and technology spokesman for the Labour Party, Jeremy Bray. In an open letter to the prime minister Margaret Thatcher, he says there is no national research programme and no coherent participation in international programmes. But Alan Apling of the Department of the Environment says that only now is the science involved mature enough to facilitate co-ordination and this is beginning to happen very rapidly. Carson says there could be more national co-ordination, and he is working to that end, producing a strategy document for atmospheric sciences in Britain to be published next year. Andrew Gilchrist, director of research in the Meteorological Office, where the country's effort in global climatological modelling is concentrated, agrees that international links in this area are not adequate.

The prime minister is expected to re-

spond formally to Bray's criticisms soon. Meanwhile, the announcement this week of a Christmas present to the NERC of £1.05 million may quieten for the moment those who claim that the government's voiced concern for the environment is not reflected in its actions. A substantial part of that money is expected to be directed to research on the greenhouse effect. And more resources for research in this area could be announced in the next few weeks when the government decides how to spend the extra money in next year's science budget. A major complaint of researchers studying the problems of man-made climate change is not one of money as much as one of staff shortages. Both at NERC and at the Meteorological Office there are difficulties in attracting good quality staff; some say this is a result of low salaries. But Andrew Gilchrist denies that the manpower shortage is as bad as it is made out to be. Though only the equivalent of one and a half people to work full time on climate modelling, there are enough back-up staff to keep the work at an international competitive standard. Researchers at the Meteorological Office do say though that work there is hindered by a shortage of computer time, and equipment to process the data the models produce. With the greenhouse effect now high on the list of government priorities, researchers now expect to receive more funds.

Christine McGourty

Dingell castigates NIH over cut grants

Washington

THE US National Institutes of Health (NIH) have been heavily criticized for taking "drastic action based solely on newspaper articles" on a case of alleged misconduct at Harvard Medical School when proper consideration was called for.

This charge is made in a strongly worded letter from Congressman John Dingell to Otis Bowen, Secretary of Health and Human Services. Dingell, who heads the Energy and Commerce subcommittee on oversight and investigations, demands an immediate explanation why NIH decided to terminate two researchers' grants before they had conducted an investigation and when they had no evidence of wrongdoing beyond "two *Boston Globe* newspaper articles".

The case concerns two Harvard researchers, Scheffer Tseng and Kenneth Kenyon, who are alleged to have had business interests in an eye ointment they were responsible for testing and to have delayed publication of results showing the ointment to be ineffective (see *Nature* 336,

506; 6 December 1988).

Both researchers lost grants but, according to Dingell's letter, an unnamed high-level NIH official admitted that there had been no "NIH investigation ... no review of the extensive supporting documentation, no discussion of the Harvard interviews of the principals, and no NIH interviews of the principals". At the time the grants were cut, the NIH official had apparently not seen the reports from Harvard or from the hospital where the research had taken place, even though six weeks had passed since newspaper reports began to appear. Dingell's committee had earlier castigated NIH for its tardy investigations of scientific misconduct. Two congressional inquiries were held on scientific misconduct in April this year.

But Dingell writes that "summary action based purely on newspaper articles was not what we had in mind" in asking the Department of Health and Human Services to develop a "rapid, effective and fair" procedure for dealing with cases of alleged misconduct.

Alun Anderson

Belgian Pasteur institute in "catastrophic" financial crisis

Brussels

THE Institut Pasteur of Brabant (IPB) is fighting for its survival because its only patron, the provincial government of Brabant, has run into a financial crisis described as "catastrophic". But a petition carrying 5,000 signatures, including 100 of scientists from abroad, has so far failed to persuade the Belgian state government to shoulder Brabant's responsibilities.

Although IPB carries the name of Pasteur, it is independent of the French Institut Pasteur at Paris and Lille. It was set up in 1900 by the Nobel prizewinner Jules Bordet, who took Pasteur's name with the permission of his widow, then still alive, and was given financial support from the provincial government of Brabant in 1903.

The institute's charter covers rabies prevention, the production of sera and vaccines, the diagnosis of infectious diseases and related research. The present staff numbers just over 125, of whom 15 are professional scientists, but there has

been no recruitment for the past ten years and the policy of not replacing members of the staff when they leave or retire is to continue.

Problems have mounted at the institute since 1967, when work began on buildings to replace the institute's antiquated laboratories to meet the then new stringency of the pharmaceutical industry. But facilities to replace the production laboratories are still not complete, a vital World Health Organisation (WHO) licence was lost and vaccine production has now ceased.

Now, the provincial government says that building work will be abandoned and the production department closed, ostensibly because it is "no longer profitable". Although profits from vaccine and reagent sales, if any, are paid over to the government of Brabant, the institute says that it has never been stipulated that production should make money, but that it will not be possible for a reapplication for the WHO licence to succeed if the new

building is not completed.

Meanwhile, the rump of the staff is under further pressure. The overall budget has been reduced by a third in the past two years, while this year's equipment budget (at BFr 2 million) is an eighth of what it was four years ago.

While accepting that IPB's vaccine production is probably too small to be economic, the institute believes that its hope that the Belgian government would allow it to remain intact by replacing Brabant's financial support has been disappointed by the national policy of regionalization aimed at defusing Belgium's long-standing linguistic troubles.

Deputy-director Professor Jean Content, whose research on interferon has won him international respect and which is now supported from national science foundation grants, says that the ending of vaccine production signals the "beginning of the end" for the institute, but that the protest over several months may have ensured that there "will be no sudden changes, but rather a slow death". He believes it ironical that Belgium should be passing up such an opportunity when Brussels sees itself becoming the capital of Europe in 1992.

Peter Coles

Natural History Museum to build DNA database in London

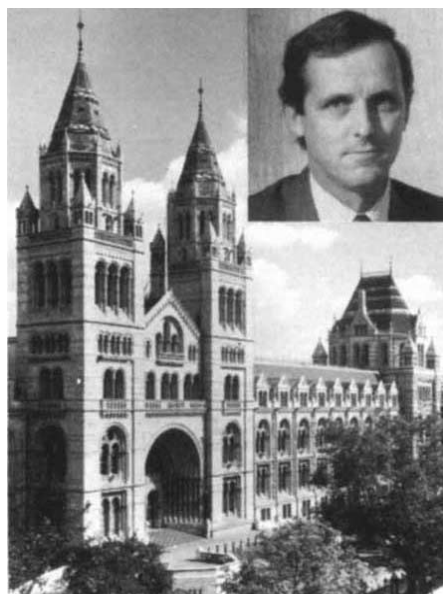
London

A BADLY needed injection of confidence among researchers at the British Museum (Natural History) in London could be on the way following the appointment last month of zoologist Neil Chalmers as director. He has ambitious plans for the museum's DNA sequencing laboratory, due to start operating in mid-1989. It will be used to extract genetic information from the museum's collections of zoological and botanical specimens to build up a large DNA database for taxonomic and phylogenetic purposes. The keeper of zoology, John Peake, has high hopes for the £200,000 facility, which will have a staff of three. It will complement biomedical work under way at two other museum laboratories to provide a "cutting edge" in taxonomic research, he says.

Chalmers answers criticisms that the DNA facility is being promoted (and funded) at the expense of more traditional lines of research by pointing out the usefulness of on-site molecular biology expertise in drawing the attention of the museum to a wider circle of biologists than it does at present. Following a review of policy at the museum, in-house researchers will be encouraged to collaborate with colleagues in other institutions, if only as an exercise to break down barriers between museum and university research. Work on extracting DNA from specimens preserved in alcohol is being done in collaboration with the Department of Genetics at the Univer-

sity of Nottingham and is an integral part of the DNA project.

And with the approach of a European free market in 1992, Chalmers has already been discussing the coordination of work in natural history museums throughout the EEC with his opposite number in Paris. Contacts with museums in other EEC



Chalmers — putting science before sentiment?

countries such as the Federal Republic of Germany are in prospect. The most visible fruits of collaboration are likely to be temporary exhibitions. An exhibition on human

palaeontology in Europe, scheduled for 1992, will be mounted jointly with a museum in Europe before it goes on tour around the EEC. Chalmers appreciates the publicity value of temporary exhibitions such as the hugely popular display of Chinese dinosaurs (to end next month), a view endorsed by the keeper of palaeontology, Robin Cocks.

But Chalmers will need to win the confidence of his research staff, demoralized by years of cuts. Although there have been no redundancies, jobs left vacant through retirements are rarely filled as resources are redistributed to new ventures, like the DNA laboratory. And worries about the long-term future of many specimens has created a sense of unease. The museum's main site in South Kensington is too small to house Europe's largest collection of natural history specimens, many of which are now stored in warehouses at Ruislip in West London. But the lease is due to expire in 1993, and no alternative site has yet been found, although the Museum is actively looking. The takeover in 1985 of the Geological Museum next door at South Kensington has freed some space, the British Geological Survey having moved to new premises at Keyworth, Nottinghamshire. Chalmers acknowledges the seriousness of the storage problem; researchers fear that valuable specimens damaged in the 8-year job of moving them to the warehouse may be damaged further when the time comes to move them again.

Henry Gee

Science and philosophy

SIR—Emile Zuckerkandl's criticism (*Nature* 334, 376; 1988) of Andrew P. Whipple's even-handed approach to science and creationism (*Nature* 333, 492; 1988) went, in my view, off the rails. The point has surprisingly been missed that the late resurgence of 'creation science' is just part of the tactics of marketing Christianity in a consumer society in which faith has been on the wane and flagging. If this point had been understood, a lot of shouting matches could have been avoided. Whipple's views are quite understandable. He argued for mutual coexistence of differing, if not contradicting, views or worldviews. The approach is a social necessity needed for peaceful coexistence. For observationally multicentred coherence, Zuckerkandl may turn to the multitude of biological researches which cannot obviously be listed in the Correspondence column. The two most important points on the issue are: (1) Science is the only course by which we know whatever we know so far about the mysteries of life; it is a process of relentless quest and the search may never end. (2) Any discrepancy or deficiency in any argument against 'creation science' neither strengthens the creationist stand nor weakens the science because creationism is unable to evolve with the progress of knowledge while science is always evolving — any weakness is the arguer's.

However, when Whipple said "there is no reality beyond the physical. . .", he spoke of physical or scientific realism. Zuckerkandl's "third position" is philosophy, not science. His assertion that "the spiritual domain has a reality of its own" is acceptable but as metaphysical or philosophical realism — surrealism, perhaps. Mixing science and philosophy would be as undesirable a hotchpotch as 'creation science' itself, which is a misnomer coined to attach credibility to a non-science. Philosophy, the rational and critical inquiry into basic principles of things, is distinct from science, the objectively verifiable sense perception/experience, although admittedly a clear line of demarcation may not always be possible. A simple example may illustrate my point. The idea that food contains vital minerals which is why we live and grow was philosophy; the isolation of those vital minerals (vitamins) and demonstration that they help us live and grow was science. Philosophy played in many instances a motherhood role for science.

Zuckerkandl cited 'mind' as an example, which is real but distinct from observable nature. Is not the concept of mind philosophical? By arguing for acknowledging the existence of a philosophical positions, he finally made it clear that he was talking from a philosophical platform. In the

dawn of human knowledge, philosophy, science, religion, superstition, astrology, astronomy, medicine and so on were all mixed up. We have left those days behind and are looking forward to the twenty-first century. We should be able to recognize the periphery of a certain discipline even though it may often be blurred, instead of standing in the blurred area and creating a smokescreen.

AZIZ ISLAM

60 Wilber Street,
Rossmoyne,
Western Australia, 6155

Use of animals

SIR—In the United Kingdom, the Animals (Scientific Procedures) Act 1986 imposes on researchers a duty to justify the use of animals in a project and then to show that the pain or distress caused is the minimum necessary. Du Pont is to sell transgenic mice that carry human oncogenes and that will develop breast cancer within a few months of birth (*Nature* 336, 300; 1988). How can the sale of such mice be justified under the 1986 act when it is in the interest of both patent holders and Du Pont to sell as many as possible? For many years, nude mice and mice carrying other naturally occurring genetic abnormalities have been sold, but genetic engineering provides for infinite expansion of the number and type of mutants. Selling such mutants does not seem to be the way to reduce the number of experimental animals used in research or to reduce the pain and distress caused.

MARK WORWOOD

Department of Haematology,
University Hospital of Wales,
Heath Park, Cardiff CF4 4XN, UK

Reprint requests

SIR—I recently published a short paper — exactly two 7 in. × 10 in. facing sides. Reprint requests began to flood in immediately, most of them from countries where photocopiers abound. None were sent out and I have had a few pointed 'second' requests. I must assume that requests were made without sight of the paper itself. The cost of sending for this reprint and the time involved are appreciably greater than walking to the library and making a single photocopy for one's own use. The cost of despatch is an expense that those whose grants do not cover the purchase and postage of reprints can well do without.

It is some years since I, and many of my colleagues, have bought reprints, being content with the free issue only. Similarly we do not request reprints unless they contain figures that greatly lose in photocopy-

ing. We can only wonder why the inveterate reprint requester has not noticed that the European reprint is a disappearing commodity. It is time that a sensible approach was taken to this problem, particularly in places where photocopiers are almost as common as journals and before scientists develop atrophy of the legs.

IVOR SMITH

Microbiology Department,
University College and Middlesex
School of Medicine,
London W1P 7LD, UK

Polio in Israel

SIR—Efforts by the Israeli health ministry to protect people in the mainland, the West Bank and Gaza against polio, which has afflicted ten victims in less than two months, have consisted of nationwide immunization of those under 40 years of age¹.

This is unfortunate, as older people have never been vaccinated with inactivated or live polio vaccine. Furthermore, isolation of poliovirus from sewage samples from seven sites¹ indicates that most older people in Israel have not had any exposure to wild poliovirus for some time and may well be negative for antibodies to three poliovirus strains. With virulent poliovirus present at at least seven sites, there is indeed a possibility of an epidemic in those aged 40 years and above, who deserve full immunization coverage — in the 1945 poliomyelitis epidemic in St Helena², there were very few cases below five years, and the maximum incidence/mortality were confined to the higher age-groups.

S.C. ARYA

Centre for Logistical
Research and Innovation,
M-122, Greater Kailash — Part II,
New Delhi-110048, India

1. *Nature* 335, 659 (1988).

2. Gear, J.H.S. in *Poliomyelitis*. World Health Organization Monograph Series No 26. 31–58 (Geneva, 1955).

Sense of proportion

SIR—Anti-abortionists lose credibility when they massively overstate legitimate concerns. Thus, when Kenneth W. M. Cochran (*Nature* 334, 560; 1988) points to the Holocaust in order to emphasize the need for governmental regulation of embryo research, he trivializes the unbearable suffering and deaths of millions caused by the Nazis (a government that certainly was strongly opposed to abortion).

To equate genocide with any lesser evil is to belittle those who suffered and to reduce the culpability of their murderers.

ARNDT VON HIPPEL

Stone Age Press,
1649 Bannister Drive,
Anchorage, Alaska 99508, USA

Using the language of science

The scientific literature, as literature, is much despised, and rightly. Should not the turning of another year be an occasion when its authors resolve to mend their ways?

THIS is the season when rash people declare their ambitions for the year ahead, sometimes recklessly in public. Secretly, a person may say to himself, or to his spouse, that he will cure some kind of cancer in 1989. It is natural, on these occasions, that there should be others who insist, self-interestedly no doubt, that the bravest ambitions are, by definition, the most likely to attain other goals as well.

What follows is a self-interested plea that contributors (and would-be contributors) to *Nature* should change their ways and, if about to cure a kind of cancer or to show the mechanism of high-temperature superconductivity, should also write so as more clearly to communicate what they have to say to at least some part of their audience. If this appeal should seem to be a dressed-up plea on behalf of the public interest, it will of course be quickly dismissed as such.

To say that present performance is appalling would be wrong; there are remarkably few complaints. Readers who write in are so much more often zealous in their regard for the proprieties of priority, or so anxious that a date should be correct, but so uncaring about the chances that their own protests will be understood, that one is bound to ask whether people who read for understanding habitually read something else instead, perhaps the back of a breakfast-cereal package.

Yet, we all believe, research has a simple structure. There is what is called a 'body of understanding' which, once understood, can be taught to undergraduates. Research is the process of adding to this structure elements of discovery which are occasionally simplifying, but which are more often complications. So do we not have an obligation to begin by reminding everybody of the starting-point, that which is already understood? "I will begin with an arresting statement that everybody can understand" should be the first of the new year's resolutions.

The second should be a resolution that the conclusion should be a declarative statement and that it should appear at least twice — at the end, where it logically belongs, and once somewhere else, preferably at the beginning. The logic is that scientific articles are different from detective stories, whose readers may enjoy suspense. Busy people, by contrast, expect that what they read will tell them in advance why they should take the trouble. And it will not for most readers suffice to

be told that "the implications for the mechanism of high-temperature superconductivity will be explored". Readers want to know what new there is to say about it.

The most serious enemy of clarity in the literature is the reverence that data command, whence a third precept. The difficulty is that any substantial programme of research is always over-flush with information; there are always numbers that cannot be quoted for lack of space, but which might be quoted if only journals were more generous. The result is that most contributions to the published literature contain more information than is necessary to support their conclusions, or even to permit mechanical replication. That is why there is the strongest possible case for not publishing information irrelevant to the intended conclusion.

But is that not an invitation to bias? If an author is encouraged not to publish all the numbers he has collected, but only those which support his case, will not the literature quickly fill with spurious arguments? That question is simple sophistry. Everybody who makes a telling point in print knows more about the real world than can be gleaned from his publication. But it is possible to tell the difference between partial and balanced arguments by the ways in which people cite their data. The strong cases are those which cite contradictory data — and then give the reasons why they do not matter. Briefly, the fourth rule is that data on their own are but numbers; they become meaningful to ordinary mortals only when their authors evidently reflect upon them.

Concepts are a further complication. The conventional wisdom is that science has been carried beyond ordinary human ken by the intricacy of its intellectual fabric. Yesterday's science may be teachable to undergraduates, but today's cannot be told in language that makes sense even to others working in the field. Can that be so? Or is it possible that the belief is a modern version of the old seventeenth-century device of concealing discoveries in anagrams or, worse, a consequence of idleness? Either way, a valuable fifth precept for intending authors would be a resolve never to make a general statement so general that readers have to cudgel their brains to know how it relates to what they know already.

Language is the next distraction. The opinion that those who practise science

cannot be expected also to have a flair for knowing how sentences are constructed and strung together is held by many, but not all, of those who earn their livings as researchers. The exceptions include many who say that there is nothing to the use of language that cannot be picked up by reading novels at weekends or even, since time is short, on airplane journeys. So, sadly, it should be plainly understood that an indifference to the way in which language is used must be equated with an indifference to meaning: a person holding that the verbs "to mitigate" and "to militate against" are in some way equivalent is no better than a person who professes not to care whether water naturally runs "uphill" or "downhill".

Naturally it is a serious matter to assert that some, perhaps many, of those who contribute to the literature of science, zealous though they may be about their numbers, are indifferent to the meaning of what they say in words. It is as if to accuse them of some acknowledged professional misdemeanour, plagiarism for example, or the deliberate misrepresentation of what data mean. But by what criteria can these misdemeanours otherwise be excused? Not to care about meaning is no better when words are in dispute. Yet the misuse of words is the most common fault in the science journals.

Sheer ugliness is evidently less grave a crime, but is so common that the question must arise of whether it is deliberate, and if so why. The innuendo is simply dealt with: ugliness is not often a subterfuge, but is more often mistaken for a mark of precision. If one molecule "binds" another, it is surely behaving much more specifically than merely "binding to" the other. The "ATP-dependent-protein kinase" cannot be just another phosphorylating agent. One obvious difficulty is that these packaged verbal constellations are often ambiguous. (Where should the hyphens go in "ATP-dependent...") A less obvious, but potentially more serious, difficulty is that many readers cannot stomach these constructions — and read something else instead.

That is why, if there are to be new year's resolutions for 1989 other than the curing of this or that kind of cancer, it would be invaluable if they could include some acknowledgement of the literary problems of the scientific literature. Why not, indeed, do both — cure cancer and let people read about it?

John Maddox

Cancer

Therapy by photon activation?

John Humm

EXPERIENCE with conventional radiotherapy has shown that to treat cancer successfully, the maximum dose possible must be delivered to the tumour — typically 50–80 Gy (5,000–8,000 rad) — compatible with the tolerances of the normal sensitive body tissues. A new approach to radiotherapy proposed by R. L. Mills *et al.* elsewhere in this issue (*Nature* **336**, 787–789; 1988), uses a Mössbauer resonance absorption technique selectively to deliver sterilizing doses to tumour cells, at the expense of only about 10^{-5} Gy to the normal tissues. There is impressive tumour ablation in tumour-bearing mice after minute radiation doses, and the authors allude to the implications for this kind of therapy in man: significant regression and maybe even cure at doses less than received by a chest X-ray.

The Mössbauer effect arises with particular radionuclide γ -transitions when the source is tightly bound within a crystal lattice. Under these conditions, the crystal as a whole can recoil, but if the recoil energy is insufficient to excite the lowest quantized phonon level of the crystal, the emission can be totally without recoil and the full emission energy is carried by the γ -photon. When this happens, a resonant reabsorption of the source emission line by an adjacent atom in the de-excited state becomes highly probable because of the exact overlap between the parent emission and the daughter absorption bandwidths. This very high-absorption cross-section, Mills and his colleagues propose, could be exploited for a novel kind of radiotherapy.

Selective absorption

Conventional radiotherapy uses high-energy γ -photons produced either by a ^{60}Co source machine or by high-energy accelerators such as a LINAC. The primary mode of energy deposition by such high-energy photons is Compton scattering, the cross-section of which varies only slightly between the different tissue constituents. If it were possible to obtain a selective absorption of radiation in some element which could be preferentially incorporated into tumour tissue via a cancer-localizing drug, then the basis of a therapy might be established. There are two potential highly selective mechanisms of photon absorption: the photoelectric effect and the Mössbauer effect.

The photoelectric effect becomes the dominant photon-interaction process in tissue at energies less than 60 keV. Because the photoelectric cross-section increases with atomic number Z as Z^4 , if a high- Z element such as iodine could be

readily and specifically incorporated into tumour tissue and then exposed to photons of energy just above the binding energy of the innermost orbital of iodine, a selective absorption in iodine might be induced when the differential cross-section of iodine compared to tissue peaks sharply. The photoelectric cross-section of iodine just above the K-shell binding energy (33.2 keV) is $38.4 \text{ cm}^2 \text{ per g}$ versus $0.15 \text{ cm}^2 \text{ per g}$ for tissue at this energy, a gain of about 250. But before too much jubilation, the predominant constituent by weight of normal tissue, oxygen, is probably at least 1,000 fold in excess of the level of iodine achievable in the cell nucleus.

But what if a high- Z element could be specifically incorporated into the DNA of the target cell? A photoelectric interaction, which induces an inner-shell electron vacancy within an atom, initiates an electronic de-excitation process involving a cascade of inner-shell transitions through the atom, the result of which is the emission of characteristic X-rays and, more important, a shower of low-energy Auger electrons. These Auger electrons, the result of many outer-shell non-radiative transitions, predominantly have ranges of only a few nanometres. The radionuclide ^{125}I , which induces inner-shell vacancies by the decay process, if appended to the DNA, the prime target for cell inactivation, can blow the DNA to pieces near the site of decay (Martin, R. F. & Haseltine, W. A. *Science* **213**, 896–898; 1981). Iodine can be selectively incorporated into the DNA of rapidly dividing tissue like tumour cells by the substitution of thymidine by an analogue iododeoxyuridine. Unfortunately, normal tissues such as the gastro-intestinal tract and bone marrow also have a rapid turnover. Proponents of photoactivation therapy therefore argue that the advantage of induced cascades by a cold analogue over a radiolabelled analogue is the ability to spare normal sensitive tissues by the beam geometry.

The excitement over the alleged potential of this technique has stimulated many groups to test its feasibility using cell cultures. The most recent study, by the group of Professor T. Ito at Tokyo University, uses Photon Factory, a synchrotron producing a tunable intense monochromatic beam of photons. The group measured the magnitude of Auger-enhanced cell killing in a mammalian cell culture with up to 20 per cent substitution of thymidine with iododeoxyuridine irradiated just above the K-edge of iodine (Shinohara, K.

et al. *Photon Factory Activity Report* **4**, 278; 1987). The effects they observed are disappointingly small: they obtain an enhancement of 1.25 in traversing the K-edge of iodine, whereas that in uniodinated cells (illustrating the degree by which the DNA is sensitized by the substituted thymidine analogue) is 2.05.

Therapeutic implications

The failure of these studies to obtain sizable enhancements that could have therapeutic implications raises the following dilemma. On the one hand, the incorporated material must be sufficient to present an interaction cross-section large enough to outweigh the effect of the dose from the secondary electron flux produced by photoelectric interactions in all the elements normally present in tissue. On the other, it must not be so high as to impair severely the fidelity of the DNA.

The work by Mills *et al.* reported in this issue at first sight offers new hope for photoactivation therapy. The highest ever measured Mössbauer cross-section is with the 14.4-keV excitation state of ^{57}Fe ($27,139 \text{ cm}^2 \text{ per g}$), about 15,000 times greater than tissue at this energy. The approach of Mills *et al.* is to use a ^{57}Fe -bleomycin conjugate. Bleomycin is a tumoricidal agent which intercalates between the strands of DNA. Mills *et al.* irradiated the tumour with a moving ^{57}Co source incorporated in a rhodium matrix. The source motion produces a Doppler shift in the γ -emission frequency and enables any mismatch between source and target environment to be retuned for resonance absorption. An incoming photon which interacts with an ^{57}Fe atom excites it to the 14.4-keV level from which it instantaneously (in about 10 ns) reverts to the ground state by internal conversion in 91 per cent of cases. Internal conversion produces an inner-shell electron vacancy and thus an Auger cascade, which could result in a high radiotoxicity for photon interactions in the ^{57}Fe in the DNA.

This study is a highly original piece of work and warrants further investigation. But I cannot help being sceptical about claims of considerable tumour regression at doses as low as 10^{-5} Gy. If the results are so great with such an insignificant dose, why not use a bigger source, deliver 0.01 Gy and cure the animals altogether? My reservations concerning the paper hinge on three points. First, a decisive control is needed, that is, moving the source at a velocity which intentionally corresponds to a Doppler shift well off the Mössbauer resonance peak. Second, the level of drug administered intra-tumorally, although well below toxicity levels for the animal, would cause considerable toxicity to the tumour, certainly if the levels of DNA intercalation reported of one drug molecule per six DNA base pairs are to be believed. On the other hand, if the drug

levels are much lower because of rapid diffusion of drug out of the tumour, then the calculated number of interactions possible in the ^{59}Fe for sufficient Mössbauer toxicity becomes too few. Finally, there is a big question as to the potential development of the technique for clinical use, and Mills *et al.* provide no indication of their views on this subject. The half-value layer of 14.4-keV photons is only 4.4 mm in tissue. An envisaged therapy could therefore be applied only to superficial lesions, for which there already exist many

other simpler irradiation procedures, for example, electron-beam, ion-beam, soft X-ray and photodynamic therapies. If the main challenge of cancer is its deep-seated and metastatic nature, then the role of photoactivation by these methods has severe limitations. □

John Humm is in the Cancer Research Campaign Laboratory, Department of Medical Oncology, Charing Cross Hospital, London W6 8RF, and in the MRC Radiobiology Unit, Chilton, Didcot, Oxford OX11 0RD, UK.

Quantum cosmology

Baby universes and making the cosmological constant zero

L. F. Abbott

CAN quantum-mechanical processes in our Universe create and destroy other universes, and if so what effects would this have on us? This intriguing question has been the subject of much recent speculation¹⁻³ culminating in the suggestion by Sidney Coleman⁴ that the creation and destruction of baby universes and the coupling of our Universe to other universes through wormholes may resolve one of particle physics' most puzzling problems, the mystery of the cosmological constant. Although this work is extremely speculative it does provide an interesting and imaginative new approach to an old and vexing problem.

The cosmological constant can be defined most simply as the energy density of the vacuum, that is, the amount of energy in a unit volume of empty space. Intuition may suggest that this quantity must be zero, but this is incorrect. In principle, the vacuum energy density can assume any value positive or negative and current ideas about particle physics and gravity suggest that it should be quite large^{5,6}. The cosmological constant can be measured by looking for the characteristic effects that a non-zero vacuum energy density would have on the geometric structure of space-time as predicted by Einstein's general relativity. No such effects have been seen and present observational limits imply that the cosmological constant is smaller than theoretical expectations by a staggering factor of about 10^{-120} . The only way to account for this enormous discrepancy between theoretical expectation and experimental reality is to assume that the parameters of nature are involved in an extraordinarily accurate and utterly mysterious conspiracy resulting in cancellations between various contributions to the vacuum energy density.

The new proposal to account for this conspiracy relies on combining the principles of quantum mechanics with general

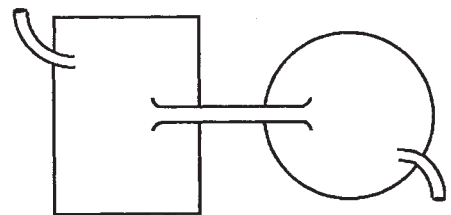
relativity, a notoriously problematic union. In fact we have no satisfactory theory of quantum gravity and even the simplest application of quantum mechanics to the geometric structure of space-time faces serious conceptual difficulties. Work on the cosmological constant is based on procedures developed by Stephen Hawking and his collaborators^{7,8} for calculating a quantum-mechanical wavefunction describing the spatial geometry of the Universe. In fact,

fluctuations. Thus they would not be directly observable. Rather, as has been shown, their effect would be transmitted indirectly through the values of the constants of nature.

The way that wormholes affect our Universe depends on the number of 'baby universes' they lead to (see figure). Rather than using the number of baby universes of type i , Coleman uses a closely related variable α_i . Because the wormholes affect the values of the parameters in physical theories and because wormhole effects are governed by the variables of α_i , all the constants of nature become functions of the α_i . This means that particle masses, the fine-structure constant, the gravitational constant and, of course, the cosmological constant all depend on parameters characterizing the topological structure of space, a remarkable concept.

Furthermore, if we can predict anything about the distribution of α_i values we may learn something about the values of physical parameters like the cosmological constant. Because the α_i s are part of our description of spatial geometry, their probability distribution is determined by the wavefunction of the Universe. Using the techniques of Hawking, Coleman finds that the probability distribution for the α_i contains a factor which is infinitely peaked at values of these parameters which make the cosmological constant vanish (for small positive cosmological

A representation of various processes involving wormholes. Our Universe is represented by the large rectangular sheet, whose horizontal direction is one spatial dimension (the other two having been ignored to simplify the diagram) and, with the qualification that is to follow, whose vertical direction can be thought of as time. The thin tubes are the wormholes. At the upper left, a wormhole appears in our Universe and creates a baby universe whose birth is represented by the circle at the top of the wormhole. In the centre, a wormhole connects our Universe to another universe depicted as a sphere. In the other spherical universe, a baby universe is being destroyed. Now for the qualification. If such processes can occur at all (and nobody knows if they can) they must be quantum-mechanical, so a simple classical space-time diagram as shown here is not applicable. But such figures are useful visual aids and if the time variable is taken to be imaginary rather than real, they provide a semi-classical approximation for the corresponding quantum process.



Hawking noted several years ago that his 'wavefunction of the universe' seemed to imply that zero is the most likely value for the cosmological constant⁹.

Coleman's new work extends Hawking's result by assuming the existence and importance of quantum fluctuations which change the topological structure of space-time (see figure). It should be stressed that we really have no idea whether such fluctuations can occur, but if they do and if their effects are relevant we can proceed to analyse what those effects might be. The first thing we know is that the connecting wormholes, filaments of distorted space-time, would have to be very tiny — about 10^{-33} cm, the natural size for gravitational quantum-mechanical

constants, it is proportional to the exponential of the exponential of one over the cosmological constant). Thus, the cosmological constant vanishes because it is infinitely more likely that the constants of nature assume values which make it vanish, than that they do not.

One of the most positive aspects of the Coleman-Hawking programme is that it avoids a major obstacle which has derailed most other attempts to adjust the cosmological constant to zero (an interesting exception is in ref. 10). Our Universe is filled with matter and radiation which throughout most of its history (presumably up to the present) have completely obscured any effects of a small but non-vanishing cosmological constant. So how

can any mechanism determine the constant's value and adjust it to zero? The answer in Coleman's approach is that the Universe peeks through a wormhole into a large empty universe thus escaping the problem of the obscuring matter and radiation in our Universe (see figure).

On the negative side, the approach relies on a shaky formalism and on many untested assumptions. It nevertheless comes up with the desired result, a zero cosmological constant. Of course, much would be forgiven if the theory could provide a correct value for another fundamental parameter, especially one that is non-zero. In principle, because Coleman's scheme is a method for predicting the values of the α_s and as all the parameters

of nature are functions of these, it is a theory of parameters. Unfortunately, early results have been disappointing so it remains to be seen whether the Coleman-Hawking approach, if it is indeed correct, will prove revolutionary or merely comforting. □

1. Hawking, S.W. *Phys. Rev. D* **37**, 904-910 (1988).
2. Lavrelashvili, G.V., Rubakov, V.A. & Tinyakov, P.-G. *JETP Lett.* **46**, 167-169 (1987).
3. Giddings, S. & Strominger, A. *Nucl. Phys. B* (in the press).
4. Coleman, S. *Nucl. Phys. B* **310**, 643-668 (1988).
5. Abbott, L.F. *Sci. Am.* **256**, 106-113 (1988).
6. Weinberg, S. *Rev. mod. Phys.* (in the press).
7. Hawking, S.W. in *Relativity, Groups and Topology II* (eds De Witt, B. & Stora, R.) 314-349 (North-Holland, 1983).
8. Hartle, J.B. & Hawking, S.W. *Phys. Rev. D* **28**, 2960 (1983).
9. Hawking, S.W. *Phys. Lett. B* **134**, 403-404 (1984).
10. Linde, A. *Phys. Lett. B* **200**, 272-274 (1988).

L.F. Abbott is in the Department of Physics, Brandeis University, Waltham, Massachusetts 02254, USA.

Sex determination

Right gene, wrong chromosome

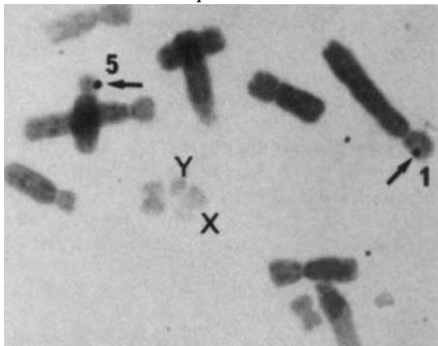
Jonathan Hodgkin

IN December last year, David Page and colleagues reported¹ on the human Y chromosome a 'zinc-finger' gene which is likely to be identical with TDF, the Y-linked testis-determining factor responsible for initiating male development. The identification of this gene, now called *ZFY*, touched off a predictable hunt for corresponding genes in other animals. In their original paper, Page *et al.* showed that *ZFY* is strongly conserved on the Y chromosome of various different placental mammals (eutherians), and also found that there is a closely related gene, *ZFX*, on the X chromosome of all these species. Now the hunt for *ZFY* cognates has extended to other vertebrate groups, with disquieting results. Exactly a year after identifying the zinc-finger gene, Sinclair, Page *et al.* report on page 780 of this issue² that the marsupial sequences most closely related to *ZFY* are on autosomes, not on the Y chromosome. This observation is based only on hybridization data, but it is unlikely to be wholly artefactual.

The result is disquieting because in marsupials, much as in eutherians, the Y chromosome is male-determining, and it is therefore expected to carry the equivalent of TDF. So if *ZFY* is autosomal in marsupials, it cannot be the primary Y-linked male determinant as it appears to be in other mammals. There are two possible conclusions: either *ZFY* is not TDF after all; or *ZFY* is testis-determining in eutherians, but something else plays the primary role in marsupials.

Several pieces of evidence favour the second possibility. The circumstantial evidence identifying *ZFY* as the human TDF is strong: for this not to be the case, it would be necessary to assume that the real TDF is an elusive, poorly conserved gene in the same small genetic interval (less

than 300 kilobases) as *ZFY*. On the other hand, *ZFY* itself is strongly conserved (at least as measured by DNA hybridization) and it also has a primary structure consistent with a regulatory role. Second, it is misleading to regard the marsupial gene or genes as corresponding to *ZFY per se*; instead they could equally or better correspond to *ZFX*. The marsupial X chromosome (see figure) is smaller than the eutherian X, and does not carry an obvious *ZFX* sequence. Several of the



Chromosomes of the kangaroo *Macropus eugenii*. (Courtesy of A.H. Sinclair.)

genes conserved on the X chromosome in all eutherians are located on autosomes in marsupials, suggesting that a translocation has taken place. Consistent with this, Sinclair *et al.*² find that a probe for the Duchenne muscular dystrophy gene, which is near *ZFX* in humans, hybridizes to chromosome 5 in wallabies, in the same interval as the apparent *ZFY* cognate.

In humans, both *ZFX* and *ZFY* could be required for testis formation. Some human XY female embryos have chromosomal defects in the Xp21 region where *ZFX* is located³, and might have a non-functional *ZFX*. Therefore, one possible explanation of the results would be that in eutherians *ZFY* potentiates the expression

or action of *ZFX*, which by itself (in XX or XO individuals) cannot trigger testis formation; whereas in marsupials some other Y-linked gene would potentiate the autosomal gene, which can be called 'ZFA'.

Genes hybridizing to a *ZFY* probe have also been detected in other vertebrate groups, but in these as well the most conserved sequences seem to be autosomal⁴. This is true of reptiles with a chromosomal sex-determination mechanism; reptiles with environmental (temperature-controlled) sex determination, such as turtles; and finally of birds, which have ZW female/ZZ male sex determination. So again, it could be that *ZFA* is consistently testis-determining, but under different primary regulation in each of these groups. Bull and co-workers have found that the turtle *ZFA* is transcribed during the critical temperature-dependent period for this species, which is consistent (but no more than that) with a role in sex determination.

Evolutionary differences in the primary sex determining signal should come as no surprise⁵. Even within a single taxonomic group such as Diptera there can be a bewildering variety of different sex-determination mechanisms, which may nevertheless turn out to have common underlying elements. Radical changes in mechanism can also be made artificially. For example, primary sex determination in the nematode *Caenorhabditis elegans* is normally achieved by the X chromosome-to-autosome ratio, as in the fruitfly *Drosophila*, yet it is possible to alter the system in various ways by mutating major autosomal sex-determining genes^{6,7}, so that the primary role is transferred to either autosome III (carrying the switch gene *tra-1*) or autosome II (carrying the switch gene *tra-2*). Many of the different vertebrate and dipteran schemes can be imitated by appropriate manipulation of nematode genes, although there seems to be little in common at the molecular level in the sex-determining genes of these three groups.

Seen from this perspective, the results obtained with *ZFY* probes are not discouraging, but they tend to focus more attention on *ZFX* as a possible major player in the process of sex determination. It remains essential to discover more about the functions of *ZFY* and *ZFX*, and what is involved in testis determination in biochemical terms: what do these zinc fingers regulate? □

1. Page, D.C. *et al.* *Cell* **51**, 1091-1104 (1987).
2. Sinclair, A.H. *et al.* *Nature* **336**, 780-783 (1988).
3. Bernstein, R. *et al.* *Science* **207**, 768-769 (1980).
4. Bull, J. J. *et al.* *Science* **242**, 567-569 (1988).
5. Bull, J. J. *Evolution of Sex-Determining Mechanisms* (Benjamin/Cummings, Menlo Park, 1983).
6. Hodgkin, J. *Nature* **304**, 267-268 (1983).
7. Miller, L.M. *et al.* *Cell* **55**, 167-183 (1988).

Jonathan Hodgkin is in the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Geochemical recycling

Why continents look too young

Paul Taylor

CAN the continental crust be recycled into the mantle whence it originally came? If transport of material from mantle to crust is essentially one-way traffic then a mean age can be determined for the continental crust using an appropriate radioactive decay scheme such as $^{147}\text{Sm} - ^{143}\text{Nd}$, notwithstanding processes of crustal reworking which may effect internal redistribution of elements within the crust. (A mean age for the continental crust gives an important constraint on models of the development of the continental crust.) But if recycling does occur then the crust is a 'leaky' system, and more complex interactions with the mantle must be considered in geochemical models. In fact, it seems that the evidence for or against recycling is highly dependent on the chemical elements under consideration. On page 733 of this issue¹, Goldstein shows that removal of radiogenic strontium from the crust causes serious underestimation of the mean age of the continents by the $^{87}\text{Rb} - ^{87}\text{Sr}$ decay scheme.

It is accepted that the buoyancy of continental (sialic) crust inhibits the bulk return of such material to the mantle. However, other mechanisms for returning some crustal constituents to the mantle have been proposed and debated. Sediments derived from the continents and deposited on the ocean floor might be carried down on the descending plate at subduction zones. Elements leached from continental rocks by weathering are carried by the rivers to the oceans in solution. Some of this material is incorporated into the oceanic crust by hydrothermal interaction at mid-ocean ridges and by low temperature alteration processes. For those elements carried in sediments on the

oceanic 'conveyor belt' (Pb, Ce, Lu, Hf, Sm and Nd), the evidence is against deep recycling into the convecting mantle^{2,3}. (Incorporation of subducted sediments into magmas for rapid return to the crust through island-arc volcanism probably serves as a filter preventing deep recycling.)

The Sm-Nd decay scheme provides a reliable basis for determining crustal residence ages because once these elements have been incorporated into the continental crust they are not lost, either by sediment subduction (see above) or by weathering and leaching (Sm and Nd are virtually insoluble in water). The Sm-Nd isotope 'clock' is based on the radioactive decay of ^{147}Sm to ^{143}Nd by α -emission. At the time of formation, continental crust acquires a markedly lower Sm/Nd ratio than that typical of its mantle source, with the consequence that the rate of change of $^{143}\text{Nd}/^{144}\text{Nd}$ ratios is significantly less in the crust than in the mantle. (^{144}Nd is a stable isotope of Nd, neither radioactive nor radiogenic. It is used as a reference isotope against which variations in the abundances of the parent ^{147}Sm and daughter ^{143}Nd isotopes can be measured.) The separation of the crust from the mantle effectively starts the isotope clock, and the crustal residence age is assessed from the contrast in the present Sm-Nd isotope compositions of the crustal sample and mantle.

Crust-mantle differentiation is apparently the main process during which fractionation of Sm-Nd ratios occurs. Subsequent crustal reworking events, whether sedimentary, igneous or metamorphic, seem to have little effect on these ratios, such that the memory of the initial crust formation age can be retained.

In his paper in this issue¹, Goldstein contrasts the behaviour of the Sm-Nd decay scheme with the related Rb-Sr decay scheme used in earlier dating studies. The Rb-Sr decay scheme gives apparent crustal residence ages which are generally much younger than Sm-Nd model ages, especially for clastic sedimentary rocks. Goldstein shows that Rb and Sr can be strongly fractionated by intra-crustal processes, not least in the sedimentary cycle of weathering-erosion-transport-deposition-diagenesis. Whereas Sm and Nd are virtually insoluble in water, weathering processes can remove Sr in solution and concentrate Rb in clays. This causes a pronounced fractionation of Rb and Sr, and removes generally radiogenic Sr from the continental crust to the ocean basins. Some of this Sr can then be loaded onto the subduction conveyor belt by hydrothermal interaction of sea water with oceanic basalts⁴, by low temperature alteration reactions on the sea floor⁵, and by deposition of pelagic carbonates.

The return of this Sr to the mantle is not directly proved by Goldstein's study, but he argues persuasively from several strands of circumstantial evidence that such return feed has taken place. First, no volumetrically significant reservoir of unsupported radiogenic Sr has been recognized within the continental crust as a complement to the material which has lost radiogenic Sr. Secondly, the systematic loss of radiogenic Sr from the continental crust can account for the near-stagnation of $^{87}\text{Sr}/^{86}\text{Sr}$ in sea water through much of the Earth's later history, and for the lowering of initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of many rock-units formed by crustal reworking. As a result of this latter effect, the Rb-Sr method results in underestimation of continental crustal residence ages, with serious consequences for some continental growth models⁶.

Finally, Goldstein also demonstrates that the return of continental Sr to the mantle could account for a very significant proportion of the total change in mantle Sr isotopic composition over the Earth's history, with serious implications for our understanding of the geochemical evolution of the mantle. This new work clearly impinges on studies of evolution of the continental crust, sea water and the upper mantle, plausibly resolving some previously enigmatic observations and placing several items of conventional wisdom under renewed scrutiny. □

1. Goldstein, S.L. *Nature* **336**, 733-738 (1988).
2. Newsom, H.E., White, W.M., Jochum, K.P. & Hofmann, A.W. *Earth planet. Sci. Lett.* **80**, 299-313 (1986).
3. Patchett, P.J., White, W.M., Feldmann, H., Kielinczuk, S. & Hofmann, A.W. *Earth planet. Sci. Lett.* **69**, 365-378 (1984).
4. Albarède, F. *et al.* *Earth planet. Sci. Lett.* **55**, 229-236 (1981).
5. Staudigel, H. & Hart, S.R. *Geochim. cosmochim. Acta* **47**, 337-350 (1983).
6. Hurley, P.M. & Rand, J.R. *Science* **164**, 1229-1242 (1969).

Paul Taylor is in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

100 years ago

THE EARLIEST RACIAL PORTRAITS

THE earliest representations of races have been strangely neglected hitherto. At the desire of the British Association I took up this work, and made a series of casts of 280 heads from the sculptures, besides noting the colours of all paintings of races that I could find.

The photograph shows the Thahennu of Northern Africa, the Kabyles of modern times.

W. M. FLINDERS PETRIE

From *Nature* **39**, 128; 6 December 1888.



Mass extinction

Of impacts and volcanoes

Laura Garwin

WHAT killed the dinosaurs? Ten years ago, this question lay firmly in the province of palaeontology. Then, in 1980, L. Alvarez *et al.* published their celebrated paper (*Science* **208**, 1095–1108; 1980) attributing the Cretaceous/Tertiary mass extinction 66 million years ago, in which 70 per cent of all species perished (including the dinosaurs), to the impact of a 10-km-diameter asteroid. The palaeontologists' monopoly of interest was broken, and a new interdisciplinary science of global catastrophes was born. A recent conference*, which spanned such diverse fields as astrophysics, palaeobotany, atmospheric chemistry and sedimentology, showed this science to be thriving.

Although there is still no consensus as to exactly what happened at the Cretaceous/Tertiary (K/T) boundary, the evidence for at least one large-body impact is now very compelling. Indeed, for many the debate has progressed to asking which of the plausible effects accompanying an impact (darkness, ozone depletion, acid rain, the greenhouse effect, tsunamis or wildfires, to name a few) may have caused the extinctions. Encouragingly, the speculations of the past have become quantitative, testable models; and field geologists and palaeontologists are adopting a corresponding rigour in describing the evidence preserved in outcrops around the world. Other mass-extinction horizons are being re-inspected for common features (see last week's News and Views article by H. Gee and R. Howlett, *Nature* **336**, 617; 1988). But it seems likely that each mass extinction arose from a distinct combination of circumstances, in which both catastrophic and gradual change were involved.

Perhaps the most compelling reason to take the impact hypothesis seriously is that asteroids and comets do strike the Earth from time to time. The Earth resides in an asteroid swarm (Fig. 1), and new Earth-crossing asteroids are still being discovered; Earth-crossing comets are even more plentiful (E.M. Shoemaker, US Geological Survey, Flagstaff). The best current estimates for the size-frequency distributions for asteroids and comets led Shoemaker to conclude that over the past 100 million years (Myr), one or two objects larger than 10 km (large enough to produce a crater more than 150 km across) have hit the Earth.

Modellers have been energetic in calculating the effects such an impact would have. Alvarez *et al.* originally proposed that dust-sized ejecta were

lofted to the stratosphere, where they could remain for several years, blocking out sunlight, halting photosynthesis and causing global cooling. J.D. O'Keefe (Caltech) suggested that the climatic effects from impact-induced dust have been over-estimated. Using the results of laboratory studies of high-velocity impacts on silicate targets, O'Keefe estimated that as little as one part in 10^6 of the ejecta would be small enough to remain in the atmosphere, and concluded that a 10-km bolide would create only just enough dust to shield the Sun. But, as the size distribution of the ejecta will be very sensitive to the target rock type (Shoemaker), this calculation cannot yet strongly

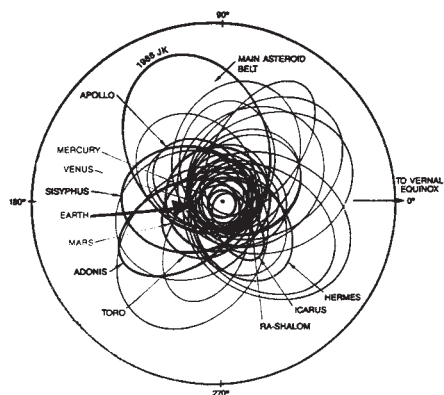


Fig. 1 Orbits of some near-Earth asteroids. About 80 Earth-crossing asteroids have been identified, and the current discovery rate of about 8 per year shows that the catalogue is far from complete, except for the brightest known objects, with diameters greater than 8 km. An asteroid of this size hits the Earth about once every 200 Myr. (Courtesy of S.J. Ostro, after an original by G.W. Wetherill.)

constrain the size of the impactor. Another complication is the soot found at K/T boundary sites around the world, which would have absorbed sunlight even more effectively than rock dust (E. Anders, Univ. Chicago).

Darkness and cooling were the two effects emphasized by Alvarez *et al.*, and one-dimensional modelling (Toon, O.B. *et al. Geol. Soc. Am. spec. Pap.* **190**, 187–200; 1982) confirmed that for up to a year light levels could have been too low for plants to photosynthesize, with continental temperatures remaining below freezing for twice this long. This study led to the concept of 'nuclear winter', but just as nuclear winter has now given way to nuclear 'fall', so new three-dimensional calculations of the atmospheric effects of an impact (P. Weissman, Jet Propulsion Lab.) yield greatly reduced temperature drops.

A three-dimensional atmospheric

circulation model incorporating NO_x formation (by atmospheric heating) and stratospheric ozone chemistry (S. L. Thompson, National Center for Atmospheric Research) gives qualitatively different effects, depending on the size of the bolide; for a 2-km bolide, greenhouse warming dominates cooling after about 10 days (but again, this depends on the assumed dust size distribution), whereas for a 10-km bolide, cooling caused by atmospheric NO_x dominates. Quantitative removal of atmospheric ozone would occur within 60 days of a 10-km impact, exposing survivors of the initial catastrophe to a bath of ultraviolet radiation.

Nitric acid formed in the atmosphere from NO_x will fall out as acid rain; a worst-case estimate gives rain of pH 0–1 for several years (R.G. Prinn, MIT). This would decalcify the upper ocean, leading to preferential destruction of calcareous pelagic organisms, as is observed. Acid rain can also explain the observed increase in nitrogen abundance at the K/T boundary (I. Gilmour, Univ. Chicago; S.R. Boyd, Open Univ.) and, at least qualitatively, the rapid increase in the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of sea water, as recorded in the shells of foraminifera from below and above the boundary (J. D. Macdougall, Scripps Institution of Oceanography). This last effect would arise from the ability of acidic rain to dissolve ^{87}Sr -rich continental granites.

Several speakers noted that a large-body impact is not the only way to cause a global catastrophe. The Krakatau eruption of 1883, which affected sunsets for several years, was less than one-hundredth the size (in terms of erupted volume) of the most violent eruptions known from the geological record. The most obvious product of such eruptions — volcanic ash — would have a negligible effect on long-term climate, because it does not stay in the atmosphere long enough. Much more important climatically are the volcanic gases and aerosols, such as sulphur compounds and halogens (H. Sigurdsson, Univ. Rhode Island). The sulphur aerosols increase the backscatter of sunlight from the stratosphere, causing surface cooling, and the halogens may catalyse ozone destruction. Sulphur compounds, at least, are more prevalent in hotspot-derived basaltic magmas (such as those that form the Hawaiian islands), and this type of volcano may in fact be more important in modifying the climate than the more violent, Mount St Helens type.

Although volcanoes can cause most, if not all of the environmental effects attributed to a large-body impact, there does seem to be a problem of scale. Just as for bolide impacts, there is an inverse relationship between magnitude and frequency for volcanic eruptions, with 1–2 Krakataus (erupted volume 10 km^3) per 100 years, 1–2 Tamboras (100 km^3) per

*Global Catastrophes in Earth History, Snowbird, Utah, 20–23 October 1988.

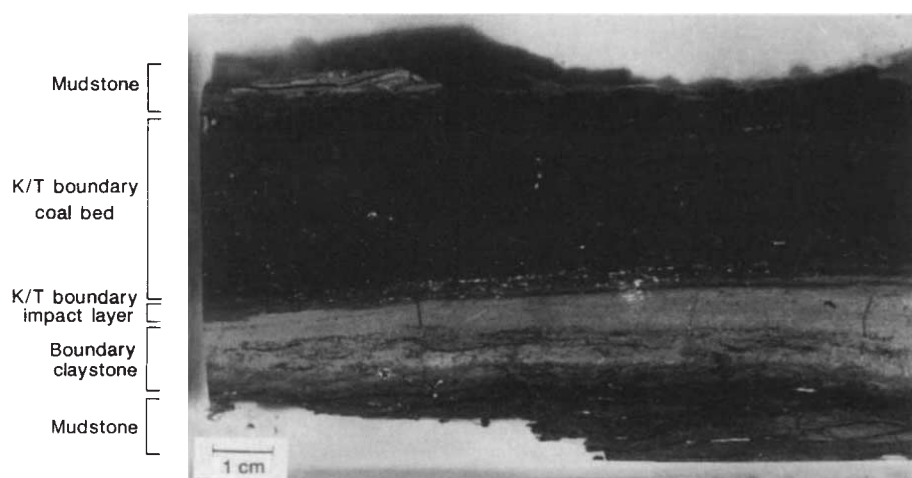


Fig. 2 The Cretaceous/Tertiary boundary interval as preserved in the Raton Basin, Colorado, showing the boundary claystone and impact layer (the latter termed the "magic layer" by G.A. Izett, owing to the sudden appearance of shocked quartz in this layer). All sites in western North America have yielded shocked grains as large as 0.5–0.6 mm, as compared with less than 0.2 mm for sites outside North America (Izett, preprint), suggesting that at least one K/T impact occurred in North America. Bruce Bohor (US Geological Survey, Denver) reported the discovery of shocked quartz in the boundary claystone, and suggested that the entire boundary layer (claystone plus magic layer) is a distal ejecta deposit. If this is so, the greater thickness of the boundary layer in western North America than elsewhere is further evidence for an impact in or near North America.

1,000 years, and 1–2 Fish Canyon Tuffs (3,000 km³) per 10,000 years (P.W. Lipman, US Geological Survey, Denver; T. Simkin, Smithsonian Institution). But whereas we know that 10-km asteroids and comets exist, and must strike the Earth, it is far from certain that sufficiently large volcanoes have existed. Indeed, the nature of the eruption mechanism of violent, ash-fall volcanoes seems to place an upper limit of about 4,300 km³ on the erupted volume (Sigurdsson). The largest flood basalt provinces produced as much as a million cubic kilometres of magma, but in the best-known of these provinces, the Columbia River basalts, the largest individual flows were only about 1,000 km³. Although these were each erupted in a week or two, they were separated by, on average, 14,000 years of quiet: it is thus not clear that the environmental effects would be cumulative (D.A. Swanson, Cascades Volcano Observatory; see also the recent News and Views article by K. G. Cox, *Nature* **333**, 802; 1988).

Of course, the ability of volcanism to cause mass extinctions is a distinct issue from that of what happened at the K/T boundary. Although many argued that volcanoes could produce the observed enrichments in iridium and other platinum-group elements, the presence of quartz grains with multiple sets of shock lamellae remains a problem for proponents of volcanism, as multiple intersecting lamellae have never been observed in volcanic quartz (Lipman; G.A. Izett, US Geological Survey, Denver). One advocate of volcanism (A.R. Huffman, Texas A&M Univ.) volunteered that the discovery of coesite or stishovite (high-pressure forms of silica — the latter known in nature only from impact sites) would, in

his opinion, "rule out volcanism"; this must have cheered J.F. McHone (Arizona State Univ.), who nevertheless waited two more days to announce that he had identified stishovite in rocks from the K/T boundary section at Raton Pass in New Mexico.

Both impact and volcanism proponents could be right, in that a large-body impact might trigger an episode of volcanism. The consensus among volcanologists was that such triggering would be inevitable, although O'Keefe deduced from his models that volcanism would occur only as the bolide penetrated downward, and would therefore be relatively weak. Shoemaker argued that a 150-km crater in the ocean would bring mantle rocks to the surface, making volcanism inevitable, but H. J. Melosh (Univ. Arizona) pointed out that the amount of pressure-release melting is controlled by the excess temperature, which for rocks from 20 km depth would be very small.

For sceptics seeking signs of an impact crater, an ⁴⁰Ar/³⁹Ar plateau age of 66 Myr for microcline feldspar recovered from the central uplift of the Manson impact structure (J.B. Hartung, Solar System Associates) makes this 35-km crater a strong candidate. Its location in Iowa ties in well with the much greater abundance and larger grain size of shocked minerals in the North American K/T boundary sites, compared with those elsewhere in the world. In addition, the Manson target rocks are mineralogically similar to the K/T shocked minerals (Izett). Although 35 km is much smaller than the canonical K/T impact crater, Izett pointed out that the size of the bolide is not well constrained, and with the additional effects of acid rain, global wildfires and the like, a smaller impact than that

originally envisaged by Alvarez *et al.* could do the job.

In fact, there may be rather an embarrassment of evidence for K/T impacts. The Manson structure can explain the shocked quartz, but other mineralogical and geochemical observations — for example, the discovery of clinopyroxene-bearing spherules at a new K/T section in Agost, Spain (J. Smit, Free Univ., Amsterdam), and the anomalously low rare-earth-element content of the boundary layer here and at other European sites — indicate a marine impact site. Smit proposed that there had been four impacts, one each for the shocked quartz, the clinopyroxene spherules, the 'tsunami bed' preserved at the boundary in Texas (J. Bourgeois, Univ. Washington), and a layer of goyazite (aluminium phosphate) spherules. At the other extreme, A. Hildebrand (Univ. Arizona) suggested that all of the observations could be explained by a single impact on the Nicaragua Rise, which is continental crust but lies beneath the Caribbean Sea. Gilmour and Anders also argued that the remarkable constancy at different K/T sites of the ratio of the trace elements antimony, arsenic and zinc (thought to come from the target rock) to iridium (from the impactor) implies that a uniform mixture of target rock and impactor was dispersed worldwide — an outcome that is improbable with multiple impacts on different target rock types.

If impacts cause extinctions, then multiple impacts extending over a period of a few million years (as in a comet shower) should create a 'stepwise' pattern of extinctions, with a succession of sudden extinction horizons (Hut, P. *et al.* *Nature* **329**, 118–126; 1987). Smit reported no evidence for gradual or stepwise extinctions of planktonic foraminifera at what he considered to be some of the most complete and expanded K/T sections in Spain, Tunisia and Israel. Yet G. Keller (Princeton Univ.) deduced from one of the same sections (El Kef, in Tunisia) that planktonic foraminifera underwent an extended period of mass extinctions starting below and ending after the K/T boundary, from which she concluded that the K/T boundary event merely hastened the demise of fauna that were already endangered by global climatic changes. Keller suggested that this disagreement might reflect differing concepts of species between the two palaeontologists, an issue that was also raised when J.D. Archibald (San Diego State Univ.) interpreted data from eastern Montana as showing that there was no mass extinction among non-marine vertebrates at the K/T boundary.

Clearly the problems of incomplete preservation, discontinuous sedimentation, and inadequate taxonomy continue to loom large, but D. McLaren (Univ. Ottawa) emphasized the importance of

being able to "sum probabilities" across many different sections, to increase one's confidence that a purported extinction pattern is real, and not merely an artefact of local circumstances. In this regard, it is encouraging that good new K/T sections (such as the one at Agost described by Smit) are still being found.

An equally broad spectrum of palaeontological views was expressed about other mass extinction horizons discussed at the meeting, including the late Eocene (36–39 Myr ago), where there is clear evidence (in the form of tektites, spherules, shocked minerals and iridium anomalies) for at least two or three impacts (B.P. Glass, Univ. Delaware; A. Montanari, Univ. California, Berkeley). An iridium anomaly is associated with the spherules, but not with the tektite layer, demonstrating — as was already suspected from studies of meteorites — that only some impactors contain enough iridium to produce an anomaly. This, combined with data from the Cenomanian/Turonian boundary (91 Myr ago), showing that an iridium anomaly comparable to that seen in the late Eocene, and about one-twentieth the size of the K/T anomaly, can apparently be produced by terrestrial processes (C.J. Orth, Los Alamos National Lab.), leads to the ironic conclusion that the evidence that first led Alvarez *et al.* to the impact hypothesis may in fact be the least reliable indicator of an impact.

Perhaps the most extreme view advanced at the meeting came from D. Raup (Univ. Chicago), who argued that all extinctions can be explained by impacts. Using the present-day impact flux and a plausible 'kill curve' relating the percentage of genera killed to bolide diameter, he succeeded in simulating the broad features of the extinction record for the past 600 Myr. He further suggested that endogenous mechanisms such as disease, predation, competition, climate deterioration and sea-level change may be effective at a local scale, but have never been shown to work on the scale of a large region or continent, at least for well established species. He cited the example of the Pleistocene glaciation — a major climate disturbance with accompanying sea-level change — which he said caused essentially no extinctions. Raup admitted that his successful simulation does not prove that impacts were the dominant agent shaping the extinction record, at least in part because it starts from the assumption that the K/T and similar mass extinctions were in fact caused by large-body impact. But it does show that, once we ascribe even one extinction to an impact, the sheer inevitability of impacts, large and small, forces us to consider the consequences for the rest of the history of life on Earth. □

Laura Garwin is Physical Sciences Editor of Nature.

Intron processing

Mobile RNA catalysts

Michael R. Green

OVER the past six years, RNA has been revealed as a catalyst of truly astonishing versatility. We now know that four distinct RNA processing reactions are naturally catalysed by RNA enzymes (see figure). The diversity of these reactions raises questions about the chemistry of RNA catalysis and the basis of substrate specificity, and has revitalized interest in the evolution and origin of introns, which in several instances are self-splicing ribozymes. Progress on these questions, reported at a recent meeting*, is selectively reviewed here.

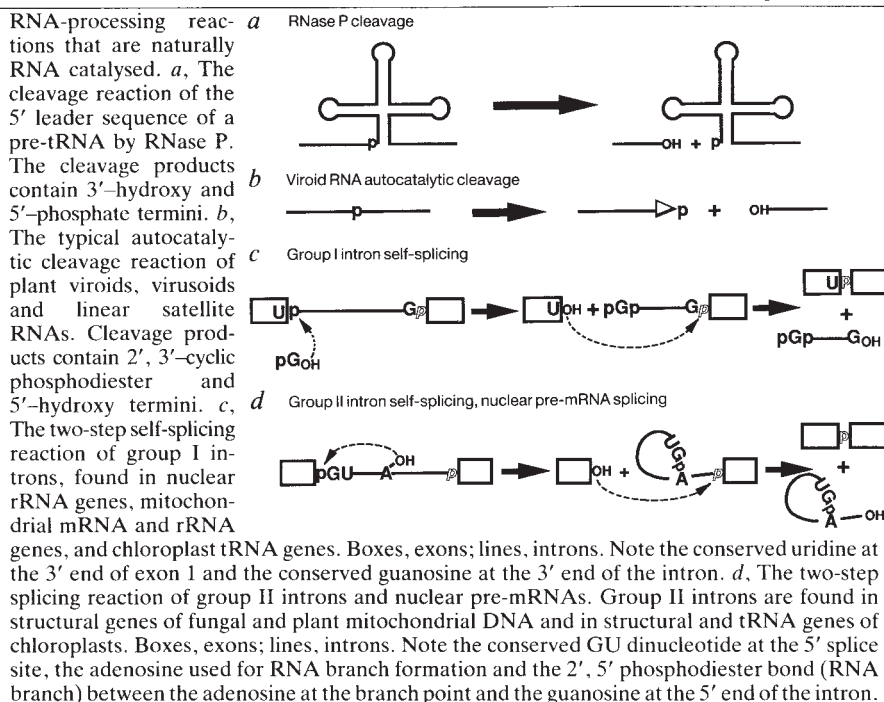
RNA cleavage reactions

The ribonucleoprotein ribonuclease P (RNase P) cleaves 5' precursor-specific sequences from pre-transfer RNAs (*a* in the figure). It is known that the RNA component of bacterial RNase P is sufficient for catalysis *in vitro*. A model for the secondary structure of RNase P RNA has been derived¹, based primarily on phylogenetic comparative analysis of RNase P RNAs from many organisms. A triumph of the phylogenetic approach is the construction of an abbreviated 263-nucleotide-long RNase P, composed solely of conserved structural elements, which functions efficiently (N. Pace, Indiana University, Bloomington). In the light of this highly conserved structure, the results of mutational studies are

somewhat surprising: extensive regions (and perhaps every residue) of the enzyme can be separately removed without completely obliterating catalytic activity or specificity (S. Altman, Yale University; Pace). This result suggests that RNase P RNA is composed of multiple catalytic elements distributed throughout the enzyme, none of which is obligatory.

A second type of RNA-catalysed cleavage reaction occurs during the replication cycle of plant viroids, virusoids and linear satellite RNAs (see ref. 2). These RNAs undergo a self-catalysed cleavage reaction as a final step in a rolling-circle RNA replication mechanism. In contrast to the potentially complex catalytic domain(s) of RNase P RNA, the catalytic domains of viroid RNAs are well defined and remarkably small. Extensive deletion analysis of the satellite RNA of tobacco ringspot virus, for example, delineates a 52-nucleotide region required for recognition and intramolecular cleavage of a 10-nucleotide RNA sequence (G. Bruening, University of California, Davis). The self-cleavage domain can potentially form a secondary structure, dubbed³ a hammerhead. On the basis of this predicted structure, small ribozymes have been synthesized that can cleave a second RNA in an intermolecular reaction⁴. Extensive mutational analysis of one such 19-nucleotide ribozyme, and a corresponding 24-nucleotide RNA substrate, has enabled a consensus hammerhead sequence to be

* 1988 Albany conference RNA: Catalysis, Splicing, Evolution, Rensselaerville, 22–25 September 1988.



derived (O. Uhlenbeck, University of Colorado, Boulder). Remarkably, only seven specific nucleotides of the ribozyme and six of the RNA substrate are absolutely required for activity. The simplicity of the active site of these ribozymes makes them extremely attractive candidates for detailed structural analysis.

Hepatitis delta virus (HDV), a circular satellite RNA of hepatitis B virus, increases the severity of viral hepatitis. HDV RNA can undergo autocatalytic cleavage and ligation *in vitro* in a manner similar to plant viroid and satellite RNAs (D. Gottlieb, Drexel University). However, the structure at the cleavage site does not resemble the hammerhead⁵, suggesting that there are undiscovered types of catalytic RNA structures waiting to be defined.

Splicing reactions

Whereas the processing reactions mentioned above involve only cleavage of the RNA substrates, splicing entails both cleavage and ligation. The essential similarities between three types of splicing reactions — group I intron self-splicing; group II intron self-splicing; and nuclear-pre-messenger RNA splicing (*c*, *d* in the figure; see ref. 6) — were enumerated by T. Cech (University of Colorado, Boulder). First, all are two-step reactions, in which the 5' splice site is cleaved in the first step and the 3' splice site cleaved in the second. Second, all the cleavage-ligation reactions seem to occur by transesterification mechanisms. Third, there are similarities in the conserved bases at the splice junctions, which when mutated can lead to comparable effects. For example, 5' splice-site mutations can result in activation of cryptic 5' splice sites in group I introns and nuclear pre-mRNAs. Conversely, mutation of the conserved guanosine at the 3' splice site abolishes 3' splice-site cleavage but not the first step of splicing for both group I introns and nuclear pre-mRNAs.

The various systems differ, however, with respect to the elements that determine selection of the 5' splice site. For nuclear pre-mRNAs, the intron sequences near the 5' splice junction are required for determining the 5' splice site, whereas in group I and group II introns, flanking exon sequences have the principal role. Single-base substitutions in intron sequences flanking a nuclear pre-mRNA 5' splice site, for example, can decrease or abolish cleavage at that site (see refs 7,8), whereas the first six nucleotides of a group II intron can be altered without affecting the accuracy of 5' splice-site cleavage (P. Perlman, Ohio State University).

The most obvious difference between self-splicing and nuclear pre-mRNA splicing is that in the latter the catalyst must be assembled from several parts. The pieces (splicing factors) include multiple proteins

George Eugene Uhlenbeck (1900–1988)

GEORGE UHLENBECK, who died on 31 October, was primarily known as one of the originators of the notion of electron spin. This idea was the result of an intensive collaboration with S.A. Goudsmit, a collaboration encouraged and stimulated by Paul Ehrenfest, their professor at Leiden University. The idea of Uhlenbeck and Goudsmit was one of the two final touches that were added during 1925 to the so-called vector model of atomic structure, the other being Pauli's exclusion principle. It

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

gave a clear-cut explanation of mysterious quantum numbers that arose in the analysis of complex spectra. Two years later Pauli showed how the spin of the electron could be dealt with in non-relativistic quantum mechanics; in the relativistic theory of Dirac, the spin appears as a necessary consequence of the formalism. And the notion of spin is not confined to the electron but is a universal feature of particle physics.

Why was this important work never rewarded with a Nobel prize? Pauli received his for the exclusion principle in 1945; would it not have been logical to let a Nobel prize to Uhlenbeck and Goudsmit follow? Was the Nobel committee disturbed by the fact that, completely unknown to Uhlenbeck and Goudsmit, Ralph Kronig had had the same idea before and had

discussed it with Pauli, who discouraged him? Or did it consider that the subsequent work of these two candidates, although valuable, was not quite of Nobel class? Neither argument sounds convincing.

Uhlenbeck's later work deals with various topics. Together with Konopinsky, he developed a somewhat different version of Fermi's theory of nuclear β -decay, which revealed interesting insights, but was later abandoned. The main emphasis of his work, however, is on statistical mechanics. With E. Beth, Uhlenbeck formulated the quantum mechanics of the second virial coefficient (the second term in a power expansion of the equation of state for a non-ideal gas). And with B. Kahn — who later, during the occupation of The Netherlands, fell victim to Nazi antisemitism — he made considerable progress in the theory of condensation. Uhlenbeck moved to the United States before the outbreak of World War II and during the war he headed the theoretical division of the Massachusetts Institute of Technology radiation laboratory, where he mainly studied problems of noise.

After the war he devoted more time to teaching. A project to write a comprehensive treatise on the foundations of statistical mechanics was frustrated by the untimely death of his collaborator T.H. Berlin, but he did publish two smaller volumes based on lecture notes. He also wrote valuable papers on the equation of state, brownian motion and the kinetic theory of transport phenomena. Uhlenbeck was an outstanding lecturer and his students, at Ann Arbor, Utrecht and Rockefeller University will remember him as a wise and benevolent counsellor. H.B.G. Casimir

H. B. G. Casimir, formerly Research Director of the Philips Company at Eindhoven, lives at De Zegge 7, 5591 TT Heeze, The Netherlands.

as well as four small nuclear ribonucleoprotein particles (snRNPs), termed U1, U2, U5 and U4/U6. These pieces are assembled with the pre-mRNA into a structure called the 'spliceosome' in which the cleavage-ligation reactions occur.

The order of snRNP binding during assembly of the yeast spliceosome, reported by J. Abelson (California Institute of Technology), turns out to be identical to the mammalian pathway, although the prerequisites may be subtly different. In yeast, for example, U1 snRNP binding requires both the 5' splice site and branch point, whereas in the mammalian system only the 5' splice site is required. In both the mammalian and yeast systems, 5' splice-site cleavage requires prior binding of U1 snRNP to the 5' splice site. But aberrant splicing of 5' splice-site mutants is not completely corrected by compensatory mutations in U1 snRNA (M. Rosbash, Brandeis University; C. Guthrie, Univer-

sity of California, San Francisco). These results implicate at least one other *cis*- or *trans*-acting component in addition to yeast U1 snRNP in specifying the precise position of 5' splice-site cleavage. Selection of the 3' splice site is potentially quite complex, as at least two *cis*-acting elements (the branch point and 3' splice site) can play a role in the appropriate experimental conditions (A. Weiner, Yale University). The ability to reconstitute snRNPs from their snRNA moieties and protein components (our own work) should facilitate future studies on spliceosome assembly.

The reaction pathways of group II intron self-splicing and nuclear pre-mRNA splicing are identical (*d* in the figure), strongly suggesting that the collective RNA moieties of the snRNPs provide at least some (perhaps all) of the catalytic function. Individual domains of self-splicing group II introns, when

added separately, can reassemble to form active catalysts⁹. But proving that spliceosomal snRNAs have a catalytic role remains a major experimental challenge.

Intron mobility and evolution

The discovery of splicing immediately prompted the question of how introns originated. The two extreme (but not necessarily exclusive) solutions as reviewed by P. Sharp (Massachusetts Institute of Technology) are either that introns were present in ancestral genes, or that introns were inserted into genes that were originally continuous. Although a decisive answer has not yet been reached, the feasibility of intron insertion has now been clearly established in several diverse systems.

One potential insertion mechanism, completely RNA driven, is that of reverse splicing in which an excised intron, which contains all of the catalytic activity, inserts back into an unrelated RNA. Reverse splicing would enable the same intron to be placed into different RNAs. The feasibility of this mechanism has been illustrated *in vitro* for a group I intron (Cech). (Of course, a DNA copy of the newly assembled RNA would have to be generated and incorporated into the genome.)

A second class of insertion mechanisms, which, in contrast to reverse splicing, is DNA-based, involves intron-encoded proteins. Introns are known to encode an assortment of products including ribosomal proteins, splicing maturases and endonucleases. One of the best-understood systems is the intron-containing 21S ribosomal RNA gene of the yeast *Saccharomyces cerevisiae*¹⁰. This intron encodes an endonuclease that cleaves at a specific site in the genes of intron-lacking (ω^-) variants. The resulting double-stranded break promotes efficient conversion to intron-containing (ω^+) forms (B. Dujon, Pasteur Institute). Analogous intron transfer by gene conversion was found in several other cases, including a *Physarum* RNA intron (P. Vogt, Cornell University); a *Chlamydomonas* chloroplast rRNA intron (C. Lemieux, Université Laval); and an *S. cerevisiae* mitochondrial structural (*cox1*) gene (Perlman). Thus gene conversion seems to be a common mechanism for transfer of group I introns in eukaryotes.

A comparable intron-transfer mechanism has also been discovered in a prokaryote, the bacteriophage T4. There are group I introns in three T4 genes (*td*, *nrdb* and *sunY*)^{11,12}. Are these prokaryotic introns relics or recent acquisitions? The sporadic presence of introns among different T4 strains is more consistent with the latter possibility, which is supported by the demonstration that two T4 introns can be efficiently transferred from intron-containing to intron-lacking T-even phases (M. Belfort, New York State Department of Health). In each case,

transfer is dependent on an open reading frame in the intron. Insertion is at a specific site within the cognate gene of the intron-less strain and requires flanking exon homology? There is no obvious selective advantage for the presence of introns, and thus the intron could be an example of a neutral molecular parasite. In any case, these studies show that introns are capable of invading the genomes of modern organisms, and are therefore not necessarily vestigial elements from the past.

Are introns in T4 a prokaryotic anomaly? Apparently not, for D. Shub (State University of New York, Albany) has also found a group I intron in the DNA polymerase gene of SP01, a *Bacillus subtilis* bacteriophage. The discovery of introns in another prokaryote emphasizes the existence of group I introns in organisms that are otherwise highly genetically dissimilar. Given Belfort's demonstration of intron transfer between related T-even phages, the possibility that group I introns have been disseminated by inter-species transfer becomes irresistible speculation.

An example of what seems to be an intron that has escaped into semi-independence has also been described. Mitochondria of the fungus *Neurospora crassa* harbour double-stranded DNA plasmids, whose nucleotide sequence and genetic organization suggest that they evolved from mitochondrial group I introns¹³. A. Lambowitz (Ohio State University) presented biochemical evidence for a mitochondrial reverse-transcriptase activity, which is encoded by the plasmid itself. The reverse transcriptase synthesizes complete ($-$) strands using the full-length ($+$) strand primary transcript of the plasmid as a template. These results imply a replication scheme in which the intron-derived plasmid propagates itself via an RNA intermediate that is copied by the reverse transcriptase back into double-stranded DNA. Several features of this process suggest that it may be a primitive type of reverse transcription, ancestral to the retroviral mechanism. These results raise the possibility that group I-related *N. crassa* mitochondrial plasmids might be a missing link in the evolutionary chain that led to retroviruses. □

1. James, B.D., Olsen, G.J., Liu, J.L. & Pace, N.R. *Cell* **52**, 19 (1988).
2. Riesner, D. & Gross, H.J. *A. Rev. Biochem.* **54**, 531 (1985).
3. Forster, A.C. & Symons, R.H. *Cell* **49**, 211 (1987).
4. Uhlenbeck, O.C. *Nature* **328**, 596 (1987).
5. Sharmeen, L. *et al. J. Virol.* **62**, 2674 (1988).
6. Cech, T.R. & Bass, B.L. *A. Rev. Biochem.* **55**, 599 (1986).
7. Padgett, R.A. *et al. A. Rev. Biochem.* **55**, 1119 (1986).
8. Green, M.R. *A. Rev. Genet.* **20**, 671 (1986).
9. Jarrell, K.A. *et al. Molec. cell. Biol.* **8**, 2361 (1988).
10. Colleaux, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 6022 (1988).
11. Chu, F.K. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3049 (1984).
12. Shub, D.A. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 1151 (1988).
13. Nargang, F.E. *et al. Cell* **38**, 441 (1984).

Michael R. Green is in the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Daedalus

Winter's icy grip

WHILE dicing with death on an icy motorway recently, Daedalus began to muse on the small amount of water necessary to destroy the adhesion of a tyre on the road. For it is, of course, the tiny film of melt-water between the ice and the tyre which does the damage: solid ice itself would give an excellent grip.

The current answer is simply to melt all the ice, by covering the whole road with a liberal dressing of corrosive salt. But Daedalus recalls the ingenious technology used on the old steam locomotives. When their wheels threatened to slip on the rails, a steam-injector sprayed sand into the exact point of contact between wheel and rail, and restored the grip. So DREADCO's chemists are bringing this idea up to date. They are experimenting with the polyacrylate 'super-slurper' compounds which soak up urine in babies' nappies. A good water-absorbing polymer can take up dozens of times its own weight of water, and still remain dry to the touch. Sprayed continuously onto a tyre, even a very light dusting of such a water absorber should take up the few micrometres of water film which do all the damage. It will be converted into a dry or slightly tacky film with excellent tractive adhesion to the solid ice beneath.

Daedalus calculates that a mere ten grams per kilometre of his 'Tyre Talcum' will keep a car utterly steady on the most treacherous ice. And as each passing vehicle adds its tiny contribution, the ice itself will slowly acquire such a loading of Tyre Talcum that it will cease to be slippery anyway: any melt-water will be absorbed as fast as it forms. Even when the ice begins to thaw, the resulting slush will be thick and doughy, with a firm, sticky grip on the passing tyres.

Thus the menace of the icy road will be overcome. The vast global investment in snow ploughs, gritting lorries, salt depots, tyre chains and so forth will be redundant; wheeled traffic will run reliably in all weathers. Pedestrians, however, will still be at risk. For their benefit Daedalus is devising a Tyre-Talcum-dispensing boot. At each step, the compression of a small air bulb sprays a light dusting of Tyre Talcum into the footprint area. Farmers, rural postmen and tottery pensioners, not to mention arctic explorers and mountaineers, should all welcome it.

But like every technological development, Tyre Talcum has its potential for social damage. Nations like Britain, whose winter is too unreliable to give its winter-sports teams the practice they need, may welcome the chance to dust it in secret and undetectable patches on ski slopes, sleigh runs, or ice hockey rinks, to sabotage their more accomplished rivals. David Jones

Tryptophans in *myb* proteins

SIR—Biedenkapp *et al.*¹ suggest that the amino terminus of *myb* proteins represents a novel DNA-binding domain structure as it bears no sequence resemblance to known DNA-binding motifs. We would like to speculate on a specific role for the periodic tryptophans in this domain.

The *v-myb* oncogene of avian myeloblastosis virus is a truncated and mutated version of its *c-myb* proto-oncogene homologue², and genes homologous to *c-myb* are found in vertebrates and in the fruitfly *Drosophila melanogaster*³. Two regions of the amino-acid sequence of *myb* proteins are highly conserved between vertebrates and *Drosophila* but only the more amino-terminal of the two is present in *v-myb*. This amino-terminal region has DNA-binding activity^{2,3} and in *c-myb* proteins consists of three imperfect repeats⁴. The cellular role of the *myb* proteins is unknown but they may be involved in transcriptional regulation.

On examining the conserved 160-amino-acid DNA-binding region of the chicken and *Drosophila* proteins³, we noticed that all nine tryptophans are conserved, and within each of the three repeats there are 18 or 19 amino acids between the first and second tryptophans and 18 amino acids between the second and third.

Because of truncation, the *v-myb* protein lacks the first and second tryptophans of the first repeat. Deletion analysis of the *v-myb* DNA-binding region suggests that a protein with only the last half of the second repeat plus the third repeat can still bind DNA, but that a protein with only the third repeat loses this property⁵. Perhaps the true functional repeat is out of phase with the sequence repeat.

Further support for the significance of the conserved tryptophans comes from the finding that the regulatory *c1* locus of *Zea mays* (maize) encodes a protein with clear homology to *myb* proteins⁶. The first 114 amino acids of the *c1* protein has two repeats similar to the second and third repeats of *c-myb* proteins. Five of the six relevant tryptophans of the *c-myb* proteins are conserved in the *c1* repeats with isoleucine replacing the first tryptophan of the third *c-myb* repeat. It may be significant that there is a tryptophan 17 residues after the last tryptophan of the second repeat of *c1*.

The evolutionary conservation of these abundant, periodic tryptophans strongly suggests that they play an important role in the structure and function of *myb* (and *c1*) proteins. We speculate that these residues may be involved in either stacking interactions and/or charge transfer interactions.

Thus, *myb* may be the prototype for a new class of DNA-binding proteins. The

chicken *c-ets-1* and *c-ets-2* genes encode nuclear proteins. In a region highly conserved between the products of these genes and those of the related human *erg* genes there is a conserved triplet of tryptophans in which 17 amino acids separate the first and second tryptophans and 18 amino acids separate the second and third⁶. The degree of divergence in this region is not sufficient for the significance of this triplet to be clear. But the E26 transforming avian retrovirus contains both the *v-myb* and *v-ets* oncogenes.

IAN A. ANTON

JONATHAN FRAMPTON

*The Beatson Institute for Cancer Research,
Garscube Estate, Bearsden,
Glasgow G61 1BD, UK*

1. Biedenkapp, H. *et al.* *Nature* **335**, 835–837 (1988).
2. Klempnauer, K.-H. & Sippel, A.E. *EMBO J.* **6**, 2719–2725 (1987).
3. Peters, C.W. B. *et al.* *EMBO J.* **6**, 3085–3090 (1987).
4. Ralston, R. & Bishop, J.M. *Nature* **306**, 803–806 (1983).
5. Paz-Ares, J. *et al.* *EMBO J.* **6**, 3553–3558 (1987).
6. Boulukos, K.E. *et al.* *EMBO J.* **7**, 697–705 (1988).

Fear of flying

SIR—I would like to draw attention to the possible hazard of the substantial quantity of uranium — more than 1,000 pounds in some cases — carried by many US-built civilian aircraft as counterweights.

Uranium counterweights are used in control surfaces of the aircraft¹, including the rudder and elevators, where space is very limited and very dense material, traditionally tungsten, must be used. The uranium used is 'depleted uranium', which contains only 0.2% of the fission isotope ²³⁵U, compared with 0.7% in natural uranium and about 3% in uranium enriched for reactors; ²³⁸U, which is not a fission isotope, constitutes the remainder.

Depleted uranium, a by-product of enriching natural uranium for use in reactors, was being sold by the US government, having 300,000 tons of it on hand, for only \$2.50 a pound in 1980. Its low price and high density led to its use not only in counterweights but also as ballast, in X-ray shielding and, particularly, as ammunition that will penetrate armour. The civilian use of depleted uranium, however, is not without its hazards. Uranium (depleted or not) is chemically toxic, slightly radioactive and, in certain forms, is classified as a fire hazard², because it spontaneously combusts on exposure to air (which is considered as a benefit for its use as ammunition). Environmental problems have arisen from the use of depleted uranium ammunition on military firing ranges³.

In aircraft, depleted uranium is only a hazard in the event of high-temperature fires that can arise after a crash. Extensive tests by the US Navy and NASA show

that temperatures in jet aircraft fuel-pool fires can reach 1,200°C, and that values of 800–1,100°C are common⁴. Such temperatures are high enough to cause very rapid oxidation of depleted uranium^{5,6}. Indeed, they are sufficient to produce ignition and complete⁷ self-sustained oxidative combustion of 8.5-mm cubes of uranium. For larger pieces, complete combustion is not inevitable; oxidation of 3-kg armour penetrators in burning tests⁸ varies from 6 to 47%.

It is the release of airborne and respirable oxide particles from such fires that presents a hazard. The extent of the release depends greatly on such factors as temperature, access to oxygen and wind speed. Largely from the penetrator burning tests, it is known that respirable (3.3-µm diameter or less) oxide particles are released in very variable amounts, with the maximum estimated at 4% of the total oxide⁸. The calculation of inhalation hazard proceeds from the maximum permissible concentration of oxide in the kidney (3 µg per gram), the weight of an adult kidney (300 g) and the fraction of inhaled uranium deposited in the kidney (2.8%). From this I calculate that about 250,000 people, at worst, could be put at risk, from the 1,000 pounds of depleted uranium in a Boeing 747.

Lower basic metal costs and the somewhat easier fabrication of large pieces seem to have been the reason for substituting depleted uranium for tungsten in US aircraft counterweights. But there is some evidence that cost was not a controlling factor. A joint National Research Council/industry report⁹, published in 1971 soon after the first 747s entered service, discussed the huge and growing stockpile of depleted uranium, and suggested that its use in aircraft would not constitute much of a safety problem. One wonders whether the decision to use depleted uranium, which was neither controversial nor visible when made nearly 20 years ago, would be made today. While the metal clearly has useful properties, any substantial commercial use would seem to require, at a minimum, that any health hazards be clearly spelled out.

ROBERT L. PARKER

*3503 1/2 Livingston Street NW,
Washington, DC 20015, USA*

1. *Metals Handbook* 9th edn Vol. 3, 773 (Am. Soc. Metals, Ohio, 1980).
2. Sax, N.I. *Dangerous Properties of Industrial Materials* 6th edn, 2711 (Van Nostrand, New York, 1984).
3. Johnson, P.K. *Int. J. refract. Hd. Met.* **2**, 179 (1983).
4. *US Naval Research Lab. Report* 5952, 122–135 (Washington, DC, 1987).
5. Baker, L. Jr & Bingle, J.D. *J. nucl. Mater.* **20**, 11 (1966).
6. Colmenares, C. A. *Progr. solid state Chem.* **9**, 139 (1975).
7. Baker, L. Jr, Schnitzlein, J.G. & Bingle, J.D. *J. nucl. Mater.* **20**, 22 (1966).
8. Mishima, J., Parkhurst, M.A., Scherpelz, R.I. & Hadlock, D.E. *Battelle Pacific Northwest Lab. Report PNL-5415* (Richland, Washington, 1985).
9. *Natn. Mat. Advisory Board Rep.* 275 (National Academy of Sciences, Washington, DC, 1971).

Prediction of solar oscillation frequencies

SIR—Deviations from spherical symmetry in the Sun cause a splitting of otherwise degenerate oscillation frequencies ν_{nlm} of modes with the same order n and degree l . This splitting is usually represented by a truncated expansion in Legendre polynomials P_k of the form

$$(\nu_{nlm} - \nu_{nl0}) = \sum_{k=1}^5 \alpha_k(L, t) P_k(m/L)$$

where t is time, $L^2 = l(l+1)$, m is the azimuthal order of the mode, and the angular brackets denote an average over n . The odd k terms result from advection by rotation; the even k terms arise from an axisymmetric aspherical component of the structure of the Sun.

Several measurements have been made of the even coefficients α_2 for acoustic oscillations (p modes), the first in December 1981–January 1982. They showed little systematic variation with L , indicating scarcely any radial variation of solar asphericity at distances from the centre greater than about 0.55 of the Sun's radius (the lower extent of the most deeply penetrating modes contributing to the 1981–82 data set, obtained at a time of high solar activity). These coefficients do, however, vary with time, and the figure shows that there is a correlation with the sunspots.

Let us suppose that a measure of the longitudinally integrated sunspot density

$$\sigma(\theta, t) = \sum_k \beta_k(t) P_k(\cos \theta)$$

is a direct indication of the acoustic asphericity in the internal structure of the Sun, in the sense that σ is proportional to the variation with respect to t and colatitude θ of the propagation speed of acoustic waves. In this case, odd terms hardly contribute to the splitting and

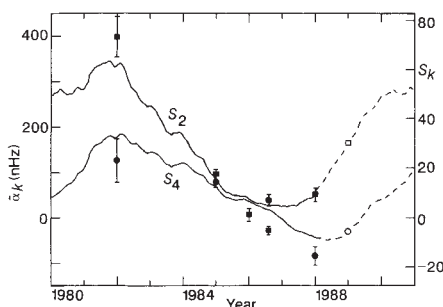
$$\alpha_{2j} \propto S_{2j} \equiv (-1)^j \frac{(2j)!}{2^{2j}(j!)^2} \beta_{2j}$$

is approximately satisfied by the even terms¹, provided by $l > j$ (which is the case for all the modes considered).

Averages $\bar{\alpha}_2$ and $\bar{\alpha}_4$ over L of the splitting coefficients α_2 and α_4 are compared with S_2 and S_4 in the figure. The direct contributions from the centrifugal distortion of the star, -22 nHz to $\bar{\alpha}_2$ and $+1$ nHz to $\bar{\alpha}_4$, have been subtracted; these were estimated from the internal angular velocity $\Omega(r, \theta)$ inferred by Brown *et al.*²

using an asymptotic frequency-splitting analysis³. The rise of $\bar{\alpha}_2$ following a sunspot minimum, announced recently⁴ and discussed in my News and Views article last week⁵, is mimicked (rather weakly) by S_2 . But S_4 lags behind S_2 , and at the time of the latest measurements it, like $\bar{\alpha}_4$, had not increased.

New observations by a team from the National Solar Observatory, United States, are being undertaken at the South Pole. If the sunspots really are a superficial indicator of an underlying acoustic asphericity, one can predict the results of those observations from the figure. The mean coefficient $\bar{\alpha}_2$ should rise further to about 140 nHz, and, following the upturn of S_4 , $\bar{\alpha}_4$ should also have started to rise,



Averaged even coefficients α_{2j} of acoustic solar oscillations. Squares, $\alpha_2 = \bar{\alpha}_2 + 22$ nHz; circles, $\alpha_4 = \bar{\alpha}_4 - 1$ nHz; data for the filled symbols before 1987–88 from ref. 6; 1987–88 data from ref. 4. Open symbols, predictions for the observations presently being carried out. Lines, S_2 and S_4 , obtained by fitting equation (2) to the longitudinally integrated relative surface density $\sigma(\theta, t)$ of sunspot numbers. (The relative density is defined such that if σ were independent of θ it would take the value of the total sunspot number over the entire surface of the Sun). Sunspot numbers obtained by summing meridian passages over intervals of one Carrington rotation period P , and then smoothing with a running mean of interval $6P$. Data from Mt Wilson archives⁷. Prediction (dashed lines) carried out using data from the previous cycle, and should be improved when sunspot data are available for the period of the current observations.

although its value, about -30 nHz, will not be substantially greater than it was last austral summer.

Finally, I consider the influence on the solar irradiance, I . If the internal acoustic asphericity is visible as a surface temperature perturbation⁶ of such magnitude that the variation of solar irradiance may be explained simply by a latitudinal redistribution of radiant flux at constant solar luminosity, then the changing shape of the asphericity during the rising phase of the cycle^{1,5} would lead to a rise of I , at the very beginning of a new cycle, which is slow compared with what one would expect if I were proportional simply to total sunspot number. The prediction from the figure would then suggest that the Solar Maximum Mission ACRIM⁷ is measuring a mean irradiance of $1,367.5 \text{ W m}^{-2}$ at present, rather than the value of $1,367.7 \text{ W m}^{-2}$ deduced on the basis of sunspot number. The irradiance variation would be roughly proportional to sunspot number during the declining phase of the cycle, however, because then the sunspots are at low

latitudes and the detailed shape of their distribution has little influence on I .

D. O. GOUGH

*Institute of Astronomy,
University of Cambridge,
Cambridge CB3 0HA, UK*

Stephan's Quintet

SIR—In his review¹ of Halton Arp's book *Quasars, Redshifts and Controversies* Michael Rowan-Robinson's only reference to a specific scientific problem is misleading.

This issue is the nature of the five closely associated galaxies called Stephan's Quintet. We showed² in 1961, well before Arp set out on his controversial career, that one of the galaxies has a much smaller redshift, by a factor of 7–8, than the other four. The question then is whether this is an accidental configuration with each component at its redshift distance, whether the difference is due to the fact that the lower redshift system NGC 7320 is being ejected from the group at a velocity of about 5,000 kilometres per second which we speculated might be the case, or whether the difference is due to some 'intrinsic' redshift property. Arp believes the last, and supposes that the four galaxies with higher redshifts have intrinsic components which dominate. Many people have studied this problem, and Arp gives an account of what he and others have done.

At the end of his account, on page 100 of the book, Arp lists the arguments for and against a large distance for the higher redshift system. Although he does editorialize, he gives a fair account. The statement by Rowan-Robinson that Arp is guilty of sophistry by giving arguments against the large distance is quite unfair. The uninitiated may well have concluded from the review that it is NGC 7319 which is the single anomalous galaxy rather than NGC 7320. To most of us, Stephan's Quintet remains a puzzle.

GEOFFREY BURBIDGE

*Center for Astrophysics and
Space Sciences,
University of California, San Diego,
La Jolla, California 92093, USA*

1. Rowan-Robinson, M. *Nature* **336**, 287 (1988).

2. Burbidge, E.M. & Burbidge, G.R. *Astrophys. J.* **134**, 244 (1961).

Seal correction

In the letter *Characterization of a seal morbillivirus* (Cosby, S.L. *et al. Nature* **336**, 115; 1988) some of the figures in the table were printed incorrectly. The results for reactivity of monoclonal antibodies against canine distemper virus (CDV) with phocine distemper virus (PDV) should read

Anti-CDV monoclonal	PDV
H1 (3.734)	—
H2 (3.755)	—
H3 (4.043)	+
H4 (2.267)	—

1. Gough, D.O. in *Seismology of the Sun and Sun-like Stars* (ed. Rolfe, E. 663–667 ESA SP-286, Noordwijk, 1988).

2. Brown, T.M. *et al. Astrophys. J.* (in the press).

3. Gough, D.O. & Taylor, P.P. *Mem. Soc. astr. Italiana*, **55**, 215–226 (1984).

4. Jefferies, S.M. *et al. Seismology of the Sun and Sun-like Stars* (ed. Rolfe, E. 277–282 ESA SP-286, Noordwijk, 1988).

5. Gough, D.O. *Nature* **336**, 618–619 (1988).

6. Kuhn, J.R. *Astrophys. J.* **331**, L131–L134 (1988).

7. Willson, R.C. & Hudson, H.S. *Nature* **332**, 810–812 (1988).

Heroes and villains

Benno Müller-Hill

Der Genetiker: Das Leben des Nikolai Timofejew-Ressowski, genannt Ur*. By Daniil Granin. Translated from the Russian by Erich Ahrndt. *Pahl-Rugenstein, Gottesweg 54, D-5000 Cologne 51: 1988. Pp. 377. Pbk DM 36.*

SINCE 1987, three editions of Daniil Granin's volume have been published in the Soviet Union. This German translation has just appeared and an English translation is underway. In the Soviet Union, public opinion is divided over the book. The conservative stalinists are horrified that a 'non-person' is honoured; some avant-gardists find the praise from Granin's mouth insufficient; and the public just loves it. The hero of this biography is a geneticist (Granin calls him *the* geneticist). What is all the fuss about?

Born in 1900, Nicolai Timoféeff-Ressovsky (or Timofejew-Ressowski) was just 18 years old when the revolution and the civil war began in the Soviet Union. He amused and charmed the anarchists who took him prisoner by telling them that Kropotkin used to come to drink tea with his father and that he, the boy, used to argue with him. In short, he survived and began to study biology. During these turbulent years much depended, perhaps even more than now, upon whom one studied with. Timoféeff-Ressovsky made the best possible choice: he went to work with Kolzov, a thoroughly well-educated zoologist who later speculated about genes being self-replicating molecules.

When Lenin died in 1924, the famous German neuroanatomist Vogt went to Moscow to procure slices from the great man's brain. There he arranged for German-Soviet scientific cooperation, and invited Soviet geneticists to come to Berlin to work in his institute. Kolzov was asked; but rather than go himself he sent Timoféeff-Ressovsky, who arrived in Berlin in 1925.

The young man had no degree and few papers to his credit, but he was very sharp. He set up a *Drosophila* laboratory to study population genetics. H.J. Muller, the American geneticist and later a Nobel prize-winner, came as a guest to the institute in late 1932 but had to leave a few months later when the Nazis came to power. Yet when Muller fled to the Soviet Union, Timoféeff-Ressovsky stayed in Germany. He stayed even when he was invited to work in the United States, for example at the Cold Spring Harbor Laboratory, and had the common sense not to go back to the Soviet Union when he heard that his former colleagues were in deep trouble. The 1930s were his

best time in science. Together with Delbrück and Zimmer, in 1935 he published the famous 'Drei-Männer-Arbeit' ('Three-Man-Paper') in which the gene became for the first time an object of interest to physicists. Schrödinger later picked up the idea, and based his famous book *What is Life?* upon it. But because Granin has no appreciation of science, we find little information about Timoféeff-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Timoféeff-Ressovsky — a survivor.

Ressovsky's actual research, besides the superlatives indicating that here was a genius at work.

The book, however, is not about science and the scientist. Rather, it is a human or social problem that Granin tries untiringly to explain: how did *the* geneticist remain sub-director of the Kaiser Wilhelm Institute until the end of the war, although he carried a Soviet passport? His answer is long-winded but simple: Timoféeff-Ressovsky was the greatest of them all, of such standing that the Nazis did not dare to touch him. His nickname in the institute 'Ur' (the bison) supposedly indicated the precious, now almost extinct class of humans he belonged to. According to Granin, he lived in his own world, so to speak, outside Nazi Germany — at least until the war against the Soviet Union. Granin says that Timoféeff-Ressovsky became aware of the persecution of the Jews in Germany only after the war began.

Life became more difficult in June 1941, when Germany launched her attack on the Soviet Union. Timoféeff-Ressovsky

began to shelter displaced European scientists, helping many endangered people by giving them safe jobs in his institute. Behind his back, his eldest son, an 18-year-old physics student, became involved in really hazardous affairs. The son worked with a group of young people who helped foreign slave workers to escape and to hide. The police got wind of the organization. The young Timoféeff-Ressovsky was arrested, jailed and then transferred to the concentration camp at Mauthausen, where he perished before the Red Army arrived. During this time his father almost broke down. He hoped until the last moment that his son would survive — to help, he stayed in Berlin even after war ended.

Despite the chaos of the first year after the war, one Soviet official was well aware that Timoféeff-Ressovsky was a formidable expert on radiation damage and wanted to get him involved in the Soviet atomic bomb project. But another official believed that he deserved to die in a work camp and arranged for him to be shipped there. It took almost a year for the benevolent official to find Timoféeff-Ressovsky, and when he was finally spotted he was on the point of death. (Solzhenitsyn has described this period, but Granin does not mention that.) After his health was restored, Timoféeff-Ressovsky was transferred to a secret laboratory. There he found his old collaborators from Germany — Zimmer, Born and others — and the work went on.

He now had to do secret research. Granin is again virtually silent about the details but apparently nothing was published during these years. Finally when the problems were solved and his German collaborators were allowed to return home to Germany, Timoféeff-Ressovsky became director of a small biological station in the Ural mountains. Although Lysenko was still in power, during the summer months he attracted students from Moscow to the laboratory, and later these people became the first Soviet molecular biologists. Timoféeff-Ressovsky was invited by the physicists to give talks on DNA in Moscow, but his enemies never gave in: he did not become a member of the Academy of Sciences and so never became part of the power élite in Soviet science.

The Soviet predilection for dividing people into heroes and villains, or saints and traitors, is fully evident in Granin's book. Forty years ago Lysenko and his colleagues were the heroes; now their enemies become the heroes. The charlatan Lysenko had been against eugenics. His former enemy, the geneticist Dubinin, was against eugenics too. But Timoféeff-Ressovsky was for it. So now when Dubinin had turned into a cantankerous old antisemite, does this not indicate that

* *The Geneticists: The Life of Nicolai Timoféeff-Ressovsky, Nicknamed the Bison.*

eugenics was and is unquestionably a good thing? But was not 'racial hygiene' another term for eugenics in Germany under the Nazis? What research did Timoféeff-Ressovsky do in this respect? Granin has nothing to say on this question. He confirms only what no one has ever doubted, that Timoféeff-Ressovsky did not do human experiments on behalf of Goebbels. Goebbels certainly had other things to worry about.

So (his research on *Drosophila* apart) what work did Timoféeff-Ressovsky publish during his years in Nazi Germany? I spent a day in the library and found the following after browsing through *Naturwissenschaften* (which regularly published the yearly record of all Kaiser Wilhelm Institutes). In 1935 he published an article on the mutational load in *Drosophila*, in which he commented that such a type of analysis would help greatly the "control" of human populations in race hygiene (*Der Erbarzt* No. 8, 117–118; 1935). In October 1938 he participated in a symposium on science and *Weltanschauung* organized by the race headquarters of the party. He invited the participants to his institute. But while he talked about mutation research, the party experts Ruttke, Gross and Wetzel talked more about *Weltanschauung*, minister Rosenberg making the concluding remarks. So although Timoféeff-Ressovsky posed as if he was laying the foundations for the race measures of Nazi Germany, in reality he did nothing. However, the public who saw his picture among all the brown shirts in the official report of the meeting (*Neues Volk* 26–30; 1939) perhaps got a different impression — in hell even a super-scientist has to pay his respect to the devils.

During the war Timoféeff-Ressovsky found an even better rationale for his research: Heisenberg's atom machine. He was certainly the best-qualified radiation expert in Germany at the time and so he extended his research in this direction. His collaborators, Gerlach, Born and Zimmer, looked at the turnover of thorium X (radium 222, an alpha emitter with a half-life of two days) in human beings (*Arch. f. exp. Pathologie* 199, 83–88; 1942). The authors of the paper do not mention who were the individuals into which they injected the thorium X, nor do they say how large the dose was. I take here their word that it was harmless. Timoféeff-Ressovsky also did not *himself* publish a single anthropological paper; again he left this to his collaborators. For example Zarapkin published a rebuttal of Boas's old and "disturbing" finding that the head form of Jews changed with the environment (*Z. Rassenkunde* 13, 113–134; 1943).

If Granin had wanted to write a true biography of Timoféeff-Ressovsky he could have easily unearthed these details, which do not diminish anything of his subject's achievements as a scientist but

which show that he was not super-human. So his book is neither a serious biography, nor a piece of art. It is just good journalism, half-truth presented convincingly. Timoféeff-Ressovsky did not deserve to die in a Soviet slave-labour camp with millions of others; it is not only saints that should have human rights. He and those other millions should be rehabilitated. Whether a serious biography of this intriguing man would be read with equal interest by the general public I doubt. But it would be well worth the effort for someone to try and write it. □

Benno Müller-Hill is a Professor in the Department of Genetics, University of Cologne, Weyertal 121, D-5000 Cologne-41, FRG.

Where and why?

Clive A. Stace

Comparative Plant Ecology: A Functional Approach to Common British Species. By J.P. Grime, J.G. Hodgson and R. Hunt. Unwin Hyman: 1988. Pp. 742. £85, \$150.

In 1941 a scheme was launched under the title *Biological Flora of the British Isles* to provide accounts of the ecological characteristics of British plants. Progress with the series, published in the *Journal of Ecology*, has been painfully slow; so far, fewer than 200 species (about 10 per cent of the native and naturalized total) have been treated, suggesting a completion date of around the year 2400! Moreover, many of the earlier accounts are already woefully incomplete and out of date, and it is clear that the scheme will never achieve its original objectives. But a synopsis amounting to an 'autecological flora' would obviously be of immense value to a wide range of people involved with any aspect of British vegetation. *Comparative Plant Ecology* is an attempt to provide the foundation for such a compendium.

The bulk of the text (562 pages) consists of 281 accounts of individual species. These accounts contain a mine of valuable and fascinating information and ensure that the book will swiftly become a standard reference work on the British flora. Each entry includes a wide range of data, among them obvious characteristics such as habitat types, altitudinal range, slope, shadiness and openness of site, and various soil characteristics, but also less expected and hence all the more welcome details such as nuclear DNA (2C) values, pollination mechanisms and a list of the five most commonly associated species. Much of the information is presented in the form of neat, easily comprehended diagrams set around a large central 'triangular ordination'.

This 'triangular ordination' is not only physically central on the page, but repres-

ents a concept central to the whole purpose and interpretation of the book. It is a depiction of the ecological strategy of each species expressed in terms of J. P. Grime's now well-known 'C-S-R model', whereby plants are considered to varying degrees as competitors, ruderals or stress-tolerators. This idea is, of course, in opposition to the even better-known concept of *r*- and *K*-strategists. Whatever one's views on the relative merits of the two approaches, *Comparative Plant Ecology* is totally wedded to the C-S-R concept, with which its users must be well acquainted if they are to interpret fully its data. Explanatory chapters give all the necessary background to enable users to reach such an understanding, as well as discussing many other important points of explanation and interpretation.

The primary data for the book have been collected over the past 25 years at the Unit of Comparative Plant Ecology at Sheffield University. Evidence of the consequent regional bias is abundant throughout the text, from the selection of species covered down to the smallest detail. It must be stressed that, although almost all the 281 species treated are certainly common (*Minuartia verna*, admittedly common around Sheffield, is a notable exception), they are not the 281 most common species nationally, and the list contains many surprises. However, further tables provide a list of 22 ecological attributes for the 281 species plus 221 others, together covering all species recorded from two-thirds or more of the 152 vice-counties of the British Isles. It can be claimed by the authors that they have presented basic ecological data for the 500 commonest British vascular plants, with the exception of maritime plants (which are excluded from the book).

As might be expected, the book is not free of errors. Most of them are trivial and merely irritating, but one old chestnut, that sycamore was prevented from reaching Britain by the formation of the English Channel (whereas in fact it does not reach the Channel on the French side as a native), should not have been repeated.

The extent to which the data will need to be modified by a wider survey remains to be seen. Revision will also be needed in certain cases because data were collected from a group of closely related species that are often confused in the field (for example *Luzula campestris/multiflora*), and where subsequent work has shown an accepted binomial to consist of distinct taxa (for example *Festuca ovina*). Nevertheless, *Comparative Plant Ecology* is an aptly titled prodromus that should be in every laboratory concerned with British plants, and one that deserves to be emulated in other parts of the world. □

Clive A. Stace is a Professor in the Department of Botany, University of Leicester, Leicester LE1 7RH, UK.

Trying to be wise before the event

Alastair Hay

In Search of Safety: Chemicals and Cancer Risk. By John D. Graham, Laura C. Green and Marc J. Roberts. *Harvard University Press*: 1988. Pp. 336. \$35, £27.95.

Living in a Chemical World: Occupational and Environmental Significance of Industrial Carcinogens. Edited by Cesare Maltoni and Irving J. Selikoff. *The New York Academy of Sciences*: 1988. Pp. 1,045. Soft cover \$260, £140.50.

SEARCHING for safety is rather like the quest for the Holy Grail. Like the Grail, absolute safety probably does not exist, but that doesn't matter when there is so much to be discovered in striving for it. And safety is intangible — it can be measured only by negatives, that is, the absence of illness, injury or dead bodies.

Sadly, we cannot talk about negatives when considering the safety record of many countries over the past few years. Take British industry and agriculture. The huge loss of life in the Piper Alpha disaster in the North Sea dwarfs every other incident recorded in annual statistics, which show that deaths from accidents on building sites and farms are on the increase, whereas they have fallen in the manufacturing and mining industries. But economic recession has forced the latter industries to contract, and reduce their workforces, and so the apparent improvement in their safety record is not real.

Accident statistics are tangible enough. More difficult to compute are the deaths caused by exposure to hazardous substances, particularly carcinogens, at work. Many of these cases never get into the records. Given the choice we would all avoid contact with compounds that might cause cancer, but for millions of workers there is no choice — exposure to these agents is a fact of life and the decisions governing them are made by someone else. *In Search of Safety* describes how some of these decisions are made in the United States, and I found the book exemplary in its discussion of the often conflicting evidence about hazards, the activities of various pressure groups and the horse-trading that is so much a part of the regulatory process.

To illustrate their theme, Graham and his co-authors chose two controversial chemicals — formaldehyde and benzene — to which we are all exposed, be it at work or through urea-formaldehyde insulation in our homes, through car exhausts and through petrol at filling stations. Rules governing both chemicals have been laid down by several regulatory

agencies and, as with most decisions in the United States, have been challenged in the courts. The challenges have often centred on the validity of the data used to determine exposures that may be safe, but which at least ought to be acceptable. In the case of formaldehyde, rulings were based primarily on evidence from experiments on animals, whereas human epidemiology served as the foundation on which to establish guidelines for benzene.

The animal evidence on formaldehyde is not clear cut. As Graham and his colleagues point out, the first problem is that although the chemical causes nasal cancer in rodents, there is only one report of a higher incidence of nasal cancer in human beings, most of the other studies describing increased incidences of lung and brain cancers, and leukaemia. Erecting a plausible hypothesis to explain how formaldehyde — which is rapidly metabolized *in vivo* — could cause cancer at sites far removed from its point of entry into the body (the lungs) has so far proved insurmountable. Equally confusing is the dose-response effect of formaldehyde on rats. At 14.3 parts per million, 50 per cent of the animals developed squamous cell carcinomas of the nasal cavity but at 5.6 parts per million only one per cent had these tumours. On the face of it, the evidence suggests that there are different carcinogenic responses operating at high and low doses, otherwise there ought to have been more tumours at the lower dose or less at the higher. These data create a problem for those who use models to predict the risk to man of low-level exposures, as there is considerable uncertainty about the number of tumours that might result. This uncertainty makes it possible for industry, or environmental groups, to challenge any

regulatory decisions in the courts.

Benzene raises other problems. Epidemiological studies in Turkey and Italy have established beyond doubt that benzene causes leukaemia in human beings, but the critical aerial concentration of the chemical is still not known. And there is some irony in the fact that regulatory agencies have been more concerned to reduce the average exposure to benzene over the working day, whereas *In Search of Safety* points out there is now good evidence to suggest that exposure to high concentrations over short periods is far more dangerous. In view of this, Graham and his co-authors suggest that "occupational exposure standards that significantly reduce allowable peak exposures might be more productive than those that strive merely to lower the TWA [time-weighted average]".

To round out the book, the authors consider the role of the courts in evaluating evidence, whether yet more scientific investigations would help, or hinder, regulatory action, and what role scientists ought to play in the process. Given the complexity of the arguments involved in regulation, they believe that the threat of litigation by industry has forced regulatory agencies to be more conservative in their rulings than the agencies would have preferred. Disapproval of this state of affairs is evident in the book: "[judges] have no special claim to wisdom about how to act in the face of uncertainty or how to make value trade-offs between economics and health in the case of particular chemicals".

On the issue of scientific evidence, Graham and his colleagues conclude that more is not always helpful, and show how in some cases additional evidence has actually confused the picture. But they do believe that more knowledge about



Visionaries — this woodcut (c. 1820) by Katsushika Hokusai shows oriental eyeglasses with lenses of tinted rock crystal. The illustration is taken from *A Spectacle of Spectacles* edited by Wolf Winkler, published by the Carl-Zeiss-Stiftung Jena/Edition Leipzig, which accompanies an exhibition currently at the National Museum of Scotland in Edinburgh, which will visit London and Liverpool in 1989. Price is £13.49.

mechanisms at the molecular level will be invaluable. As for scientists themselves, the authors consider that they need to be more humble, admit that they do not have all the answers and make that plain to the public. Science will help in the longer term. But, in the short term, negotiations aimed at reaching a consensus view are far more valuable.

Much of the information that is necessary to reach a consensus on the dangers from particular chemicals is to be found in *Living in a Chemical World*. The book is the proceedings of a conference held in Italy at the end of 1985, and the contributors address all aspects of carcinogenicity from the molecular level to epidemiology, using a whole array of chemicals to illustrate their points. The contents probably live up to the blessing conferred on the conference by Pope John Paul II, who sent his "heartfelt implored apostolic propitiatory blessing for abundant heavenly favour and enlightenment". Interpreting the blessing was far more difficult than following most of the science in the 1,000 or so pages of this book, which provide plenty of enlightenment, much as treasure is to be found along the way in searching for the Holy Grail. □

Alastair Hay is a Lecturer in the Department of Chemical Pathology, Old Medical School, University of Leeds, Leeds LS2 9JT, UK.

ADVERTISEMENT

Manipulators with piezos

Piezoelectric Manipulator PM 20 N

for fast stabbing movement. Maximum travel 20 μm . Step lengths up to 18 μm . For all intracellular manipulations. With booster up to 250 $\mu\text{m}/\text{msec}$ = 250 mm/sec! Vibrator frequency 0–100 Hz.

Step Motor Manipulator MF 500

for positioning microelectrodes in biological tissue. Precise step movement with minimal lateral vibration and no backlash at speeds up to 1 $\mu\text{m}/\text{msec}$. Total travel 45 mm!

Hybrid Piezoelectric Manipulator PM 20 H

Ideally suited for all cell manipulations. Step lengths between 0.5 and 9 μm . This is a combination of the piezoelectric device PM 20 N (see above) with a manipulator driven by motors along all three axes.

Advantage: FAST by piezo technique, LARGE TRAVEL by motor drive on three axes. Coarse adjustment by hand. Available with vibrator up to 100 Hz.

With booster: step speeds up to 250 $\mu\text{m}/\text{msec}$!

Vibrator with variable frequency available.

NEW VERSION of the apparatus position (4): PM 10 H

Simplified control, no digital display of travel, somewhat lower precision. Has piezoelectric translator like the one described in position (2), maximum travel 10 μm .

Extremely economical version for manipulations of and within cells!

BACHOFER GmbH — Reutlingen D 7410
P.O.B. 7089 — W. Germany Fax 7121-54000

Measured progress

R.M. Simmons

Molecular Mechanism of Muscle Contraction. Edited by Haruo Sugi and Gerald H. Pollack. Plenum:1988. Pp.742. \$125, £74.

THE previous volume in this series of conferences on muscular contraction was reviewed for *Nature* (311, 767; 1984) by my former colleague Mike Spencer. Under the title "Spasm? Or rigor?" he used the occasion to make some unkind (and I gather to some readers offensive) remarks about progress in understanding contraction at the molecular level. It had seemed, he wrote, in the early 1970s that a grand synthesis of structure and biochemistry was imminent, the 'muscle problem' would be solved, Nobel prizes would be in the bag. This promise had not been fulfilled, conventional wisdom was in tatters, anarchy ruled and so forth. I believe that this view is in large part demonstrably untrue and that, as a commentary on how quantitative science proceeds, it is completely misguided.

There is, however, no smoke without fire, or for that matter without actual damp. Spencer rightly appreciated that one of the main bars to progress has been the lack of definitive evidence about the nature of the structural change that underlies force generation. In an active muscle it is believed that myosin cross-bridges from the thick filaments interact cyclically with sites on the actin-containing thin filaments, each cycle involving the hydrolysis of a molecule of ATP. At some point in the cycle there is a conformational change, which in the simplest model of the early 1970s is a rotation of the head of the myosin molecule about its point of attachment to actin. We now know a great deal about the relation between the biochemical cycle and mechanical events, with steady and in some cases (such as the use of caged-ATP) spectacular progress. The theoretical formulation of general models of the cross-bridge cycle is well established, and some specific examples have been worked out. There is a long way to go with biochemistry, mechanics and theory, but conceptually the problem seems solved or solvable.

But, to return to structure, although there are now some striking correlations between structure, mechanics and biochemistry, there seem to be unresolved differences between the results of some of the techniques, or their interpretation. At the root of the problem, though, is the fact that the well-ordered cross-bridge arrangement seen in relaxed muscle and in rigor becomes disordered during an active contraction, because cross-bridge cycles are not synchronized. Further, no simple analogue of the force-generating transi-

tion exists. To interpret the changes in contraction will require a combination of some extremely painstaking electron microscopy (with fast-freezing methods now available) and time-resolved X-ray diffraction, plus optical probe and ESR techniques, plus the use of fragments of myosin with NMR and the other techniques. Protein crystallography, the use of mutant animals and site-directed mutagenesis are beginning to make an impact and will assume a greater importance as we get closer to the details of the actin-myosin interaction. With novel motility assays it is becoming possible to examine the interaction at the single molecule level. There has never been a time in the history of research on muscular contraction when such a rich diversity of techniques has been brought to bear on the problem, nor when realistically it has been possible to define a solution.

In many fields a line would have been drawn in the 1970s and we would have had to be satisfied with the then conventional wisdom. But, in biology, muscle is special in two respects. It has an unusually well-organized structure in many tissues so that it is possible to make accurate measurements. Then there is a long tradition of quantitative science in the field, so that the measurements are actually made. My reading of the past two decades is of these two aspects building up to something of a fruition in which measurement informs crude observation and theory in a way that is impossible in many fields.

Several of these structural approaches as well as biochemistry and mechanics are featured in this book, which derives from the third of a series of conferences organized by Pollack and Sugi. One similarity between this review and Spencer's is the complaint that the volume appears two years after the conference. The book has 742 pages of contributions plus discussion, and apart from one or two eccentric entries, and a bias towards the more physiological aspects of contraction, it is fairly typical of what happens at muscle contraction workshops. It is (almost) strictly for the professional, although we have found the previous volumes a useful source of topical papers for advanced undergraduates. In my experience of the biological sciences, research workers are usually either discoverers or measurers. This volume is one for the measurers. □

R.M. Simmons is a Professor in the Division of Biomolecular Sciences, King's College London, 26–29 Drury Lane, London WC2B 5RL, UK.

New in paperback

• *An Urchin in the Storm* by Stephen Jay Gould. Publisher is W.W. Norton, price is \$7.95. To be published in the UK by Collins. For review see *Nature* 330, 280 (1987).

• *Solid-State Chemistry: Techniques* edited by A.K. Cheetham and P. Day. Publisher is Oxford University Press, price is £15, \$70. For review see *Nature* 329, 115 (1987).

Is the space environment at risk?

G. B. Field, M. J. Rees and D. N. Spergel

The huge potential of space for science, exploration and peaceful applications could be jeopardized by debris in Earth orbit, and by an extraterrestrial arms race.

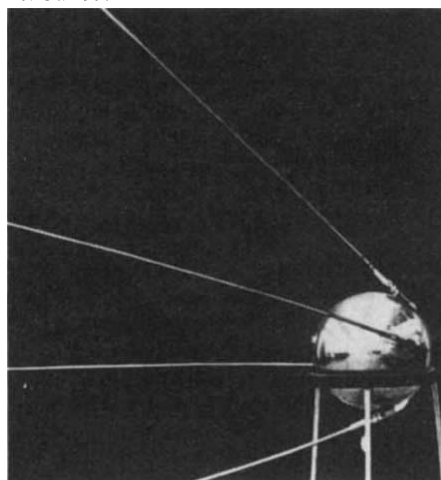
FOR many years there has been public awareness of the depredations that modern mankind has brought upon the Earth. As we approach the end of the twentieth century, similar effects beyond the planet are becoming apparent. Here we summarize the issues recently discussed by an international group of space scientists from Europe, the United States, the Soviet Union and Japan*. Our conclusion is that space-faring nations must show self-restraint in certain activities, and cooperate in others, if the hazards now evident are to be avoided and the enormous rewards offered by space-based instruments are to be realized. Much of the technology now being developed primarily for military purposes in space can be applied, when the international situation permits, to space science, global communications, navigation, environmental monitoring and arms control.

Environmental monitoring. Space vehicles offer unique advantages for monitoring changes on the Earth, on its oceans and in its atmosphere. When combined with localized ground-based measurements, a programme of observations from space yields comprehensive data about our planet on a global and daily basis. Satellites have already greatly improved weather forecasting and may be used to monitor deforestation, desertification, volcanic activity, the state of the sea surface, oceanic currents such as El Niño, and the abundance of trace gases in the upper atmosphere. The energy budget of the oceans, the solar wind, and the particles trapped in the Earth's magnetic field (all of which affect the terrestrial environment to varying degrees) can also be monitored from space.

The decrease of ozone and increases of carbon dioxide and methane in the atmosphere threaten to have profound consequences for Earth's biota. Space-based monitoring is essential for obtaining enough understanding of the causes of these changes to respond by formulating rational environmental and industrial policies. Global observations of the 'greenhouse gases', resolved regionally and in time, are necessary in order to identify the

natural and unavoidable contributions to these atmospheric gases, and the sources and relative scale of man-made additions. Similar observations of ozone and ozone-controlling trace constituents will be needed to distinguish seasonal and solar-cycle variations of natural 'pollutants' from possible depletion by man-made

"Only 30 years after the launch of Sputnik [below], the time has already come to ensure the preservation of the space around our planet as a common resource."



chemicals, and to assess the scale of other detrimental environmental changes (deforestation for example) resulting from human activities.

Nations with the appropriate facilities should accordingly establish a cooperative programme for continuous atmospheric and environmental monitoring; scientists from other nations, including the Third World, should be enabled to participate. The International Geosphere-Biosphere Program, launched by the International Council of Scientific Unions, is a welcome initiative in this regard.

Space debris. Man-made debris in orbit about the Earth is a real and growing threat to spacecraft. Collision with a single fragment a centimetre or less in size can destroy a spacecraft. Several satellites may already have been damaged or destroyed in this way. The hazard is aggravated by the fact that a single space launch can leave hundreds or thousands of highly destructive fragments in orbit. Because such debris tends to disperse, relative velocities are high while avoidance is

extremely difficult. Most fragments would remain in orbit for hundreds of years. If the density of debris becomes high enough, hypervelocity collisions will lead to further fragmentation, creating an exponentially increasing number of hazardous fragments.

If present practices continue, it may only be a decade before the problem reaches unacceptable levels. A radical change in attitude is called for if humanity is to continue to enjoy the benefits of satellites in Earth orbit. Nations should not intentionally add to the existing debris by military activities such as anti-satellite or anti-ballistic missile tests; civilian space users should take pains to reduce the risks of accidents that create debris, and should cooperate in evaluating control measures, and developing practices that alleviate the hazard.

Nuclear reactors in space. Two orbiting nuclear reactors have already re-entered the Earth's atmosphere from space; a further occurrence, involving Cosmos 1900, was narrowly averted in September this year. Even when in their proper orbits, the presence of these reactors is detrimental to some scientific programmes: gamma ray observations of celestial objects have for several years been confused and contaminated by the effects of radiation leakage from satellites.

If nuclear reactors were restricted to altitudes above 5,000 kilometres, their orbital decay times would be thousands of years and there would be much less danger of their being destroyed by space debris. On environmental grounds, nuclear reactors should be banned from low Earth orbit. Indeed other considerations relating to national security (discussed below) favour the complete prohibition of nuclear reactors in space.

Military versus peaceful purposes. During the 1960s, the huge amount of funding for the Apollo programme was accompanied by a comparable reduction in expenditures on US strategic forces. The recent increase in the scope of the Soviet civilian space programme has similarly been accompanied by reductions in the production of strategic weapons by the Soviet Union. Further cut-backs in nuclear forces would allow redirection of the specialized talents and production capacities of aerospace industries. Among the possible projects to be considered are a global environmental programme

* The Second Saltmarsh Meeting was held on 13-15 July 1988 at King's College, Cambridge, UK, with the support of the MacArthur Foundation. The other participants were M. Barnett, R. D. Blandford, R. Brandenburger, Y. Galparin, R. L. Garwin, T. Gold, M. Grubb, M.C.E. Huber, B. Jasani, F. K. Lamb, B. Morel, M. Oda, J. P. Ostriker, J. Pike, K. A. Pounds, J. Primack, G. Prins, M. Rowan-Robinson, R. Z. Sagdeev, D. Southwood and P. Zimmerman.

supported by space-based monitoring, improvement in global communications and large-scale scientific projects. International collaborations are not only valuable scientifically but provide a substitute for military build-up, thereby reducing the risk of war.

Specific technologies now being developed for military use include ultra-strong and ultra-light materials, global communication systems with large data-storage capability, advanced robotics, more efficient propulsion systems and precision optical systems (along with the means to point them at remote targets). Several of these have similar commercial potential to that already realized in satellite communications; insofar as the international situation permits, the relevant technologies should be made available to the civilian sector.

Cooperation in space science. International collaboration can involve projects of either small or large scale, with a correspondingly broad range of costs and periods of time over which the project is developed. Several relatively inexpensive and exciting collaborations could be implemented rapidly, among them the launching by Soviet boosters of small US and European scientific payloads whose launch has been delayed by setbacks to the US shuttle programme. Questions of technology transfer can be dealt with as appropriate and must not be allowed to interfere with these promising scientific opportunities.

Collaborations depend ultimately not on equipment but on people. Interchanges of staff — both junior and senior researchers, as well as students — enhance scientific and personal contacts and thereby provide the groundwork for yet further cooperative research. Such contacts would be greatly facilitated by the establishment of an International Space Institute open to scientists from all nations.

Some larger-scale projects, such as space-based radio interferometry, are intrinsically international. As experience of collaboration accumulates, the time will come for joint projects on a bigger scale, such as large orbiting astronomical observatories or sophisticated unmanned planetary missions (for example a Mars rover and sample return).

Endeavours on a still larger scale, exemplified by manned space programmes, cannot be expected to yield a return commensurate with their cost if the judgement is made in purely scientific terms. But an increased human presence in space may be regarded as worthwhile for civilization on other grounds.

Space science is valuable not only in its own right but as a driver of new technology with commercial potential. The aspiration to understand the Earth's place in the Universe creates demands for new instruments, which offer major engineering

challenges, just as military requirements have done. In meeting these challenges, industry will inevitably develop new concepts, processes and materials that can be applied elsewhere.

Monitoring of orbiting objects. At present there are more than 7,000 tracked objects in Earth orbit, of which about 350 are active payloads. The 1976 Convention on Registration of Objects Launched into Outer Space requires the registration of each payload together with its character, orbital elements and information on each launch. Over time, however, both because of natural perturbations and deliberate manoeuvring of some satellites, almost all public information on satellite orbits is lost. Each space-faring nation should provide the United Nations registry with periodic reports of the orbital elements of active satellites and, so far as is possible, of its inactive payloads as well.

Reporting such information at intervals of roughly 30 days is frequent enough to reduce the accumulation in naturally caused errors in the orbital elements of satellites. It is, on the other hand, infrequent enough to preserve the confidentiality of the timing of deliberate manoeuvres of surveillance satellites.

The accuracy of such nationally provided catalogues can be monitored at moderate cost using small optical and radio telescopes on the ground. The establishment of centres for such monitoring, both by members of the United Nations, and by private groups and non-governmental or international scientific organizations, should be encouraged.

Verification of arms control treaties. One of the most important applications of remote sensing is in the area of international security, particularly through monitoring of arms control agreements. Inability to verify compliance has often been given as a reason why such agreements are not reached. Until now, multilateral arms control treaties have been verified using only National Technical Means (that is, military 'spy' satellites). Not all nations have such means. Thus we must consider the establishment of an international verification agency, especially as civilian spacecraft are now sophisticated enough for some relevant purposes, being already able to detect objects on the Earth as small as five metres in size. Cooperation in the verification of arms control agreements could be achieved in stages, starting with feasibility studies on an 'academic' scale (as suggested at last year's Saltmarsh meeting), and developing — if the outcome of such studies were encouraging — to a multinational, regional and eventually a full international agency.

Anti-satellite capability and space weapons. In order to preserve space for scientific and commercial purposes (as well as for 'non-weapon' military uses

such as navigation and communication), and to ensure that space activities aid the limitation of strategic nuclear forces rather than encourage their expansion, there should be a total ban on the development, testing and deployment of all dedicated anti-satellite (ASAT) weapons and on space weapons of any kind. Strict observance of the 1972 anti-ballistic missile (ABM) treaty as signed and ratified, and as confirmed by the US Senate in 1987 and 1988, is equally important. The ABM treaty and the banning of anti-satellite weapons are mutually supportive, in that lack of either harms the other.

In view of the powerful ASAT capabilities of ABM systems, and in the context of multilateral restraints on strategic offensive weapons, there needs to be a further protocol to the ABM treaty reducing the number of permitted sites from one to zero. This would require the Soviet Union to dismantle its ABM system around Moscow, and that the United States should neither reactivate its system at Grand Forks, North Dakota, nor develop any new ones.

An effective treaty to guard against the proliferation of anti-satellite and space weapons must incorporate restrictions to ensure that such weapons cannot be acquired so rapidly that they could conceivably prevail against countermeasures of comparable systems. A sufficiently long 'buffer time' after abrogation of such a treaty can be achieved by banning not only the deployment but also the development and testing of these weapons. The 'buffer' could be further strengthened by a ban on all testing of nuclear reactors in space (solar energy provides adequate power for other functions of satellites in Earth orbit). Even if nuclear power for deep-space vehicles were permitted, devices should be tested in a manner open to all nations.

Conclusion. Only 30 years after the launch of Sputnik, the time has already come to ensure the preservation of the space around our planet as a common resource. There are opportunities for wider international collaborations in the monitoring of Earth's atmosphere, the verification of treaty compliance, and the pursuit of other peaceful and scientific goals. The problems are twofold: increasing contamination of space with debris, and the prospect of a dangerous and wasteful extraterrestrial arms race. It is not only the scientists of the future, but the whole of humanity, who will rightly hold the scientists and politicians of today responsible if we fail to resolve them. □

G. B. Field is in the Department of Astronomy, Harvard University, Cambridge, Massachusetts 02138, USA. M. J. Rees is in the Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA, UK. D. N. Spergel is in the Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA.

Golden jubilee for the coelacanth *Latimeria chalumnae*

Peter L. Forey

Department of Palaeontology, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK

Fifty years ago this week, Latimeria chalumnae was discovered, the only living representative of the otherwise extinct coelacanth fishes. Half a century of research shows it is not the hoped for missing link between fish and land vertebrates and there are now fears that it too will become extinct before its centenary.

ON 23 December 1938, a large steel-blue and white fish was landed at East London, South Africa. Its uniqueness was quickly spotted¹ and it was identified as a living coelacanth, one of a group of fishes, until then only known from fossils that were thought to have died out in the Late Cretaceous (about 80 million years ago). Named *Latimeria chalumnae*², the new discovery resembles fossil coelacanths and was soon dubbed a 'living fossil', a little-changed survivor of a once-abundant group. Subsequent specimens have been caught in the Comores Archipelago³, in what proves to be its natural habitat (Fig. 1).

The discovery of *Latimeria*⁴ raised hopes of gathering direct information on the transition from fish to amphibians, for there was then a long-held belief that coelacanths were close to the ancestry of tetrapods. *Latimeria* was thus heralded as a 'missing link', a reputation that rested on the theory that coelacanths and tetrapods were more closely related to each other than either was to any other living group. But studies of the anatomy and physiology of *Latimeria*⁵⁻¹³ have found this theory of relationship to be wanting and the living coelacanth's reputation as a missing link seems unjustified. On a broader front, the discovery of *Latimeria* has provided a test of the ability of palaeontologists to reconstruct fossil organisms from the often fragmentary remains in the fossil record as well as offering a cautionary reminder that we still know very little about life in the oceans, highlighting the need for further exploration and determined attempts at the conservation¹⁴ of this extraordinary creature.

An evolutionary case history

Throughout their history, coelacanths have been predominantly marine. They seem to have reached their maximum diversity between about 248 and 213 million years ago in the Triassic Period. But their morphology has always remained distinctive and seemingly conservative (Fig. 2). This helped in the rapid identification of the living fish², which is separated from its nearest fossil relative *Macropoma* by some 80 million years. As in most coelacanths (Fig. 2), the body of *Latimeria* is plump, the paired limbs, second dorsal and anal fins are lobed, the first dorsal fin is located well forward and is sail-like, the caudal fin is deep, with a small central tuft-like supplementary lobe posteriorly. The scales are large, robust and covered with many small denticles. The skeletons of *Latimeria* and *Macropoma* are practically identical, although the ossified gas bladder of *Macropoma* is replaced by a soft-walled, fat-filled organ in *Latimeria*.

Such morphological conservatism has raised questions about the tempo and mode of evolution¹⁵, and evolutionary studies of both coelacanths^{16,17} and lungfishes¹⁸ (see Box) have been used to support models of how changes in evolutionary rate may explain the histories of long-lived groups¹⁹. One scheme proposes that initially rapid morphological evolution under

intense directional selection is followed by a phase of stabilizing selection in which evolution slows down and may even stop^{19,20}. For coelacanths, but not lungfishes²¹, this pattern was claimed to be independent of changes in the numbers of species with time; the rate of morphological change was greatest in the late Devonian, some 120 million years before species diversity reached its peak (Fig. 2). But without a clear idea of the phylogeny of coelacanth genera, especially before *Latimeria* was discovered, it seems that this was based merely on morphological deviations from a theoretical ancestral type. Data from *Latimeria* allowed the construction of outline classifications of coelacanth genera^{17,22,23} justified by distributions of morphological characters. Adding a time dimension (Fig. 3) showed that earlier ideas of coelacanth evolution were mistaken: the times of maximum species diversity and the greatest rate of inferred morphological change coincide in the Triassic.

But the reasons for the slow rate of change over much of the history of the coelacanths is a separate problem which has elicited contradictory explanations. One view stems from

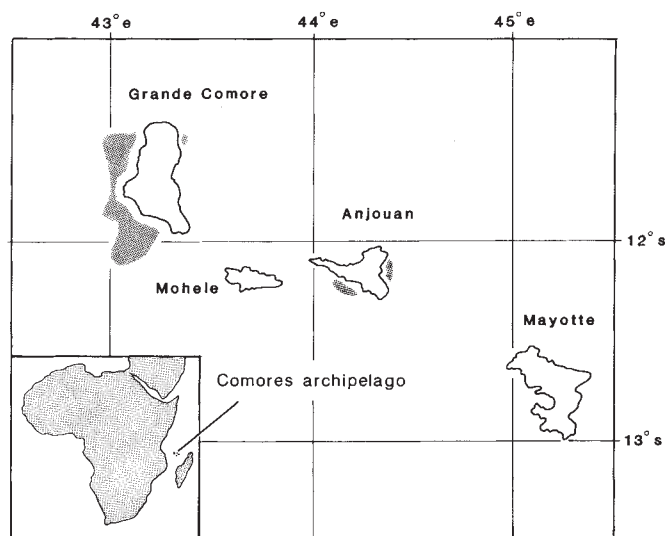


Fig. 1 *Latimeria* inhabits the steep submarine slopes off Grande Comore and Anjouan in the Comores Archipelago, between the African continent and Madagascar. Most catches^{14,25,60} have been taken from the relatively sand-free western slopes of Grande Comore at depths between 60–400 m. The fishes²⁵ prefer cool waters, 17–25 °C, and migrate diurnally. Prey items recorded include⁶⁰ cuttlefish and a variety of deep-water reef fishes. Areas where coelacanths have been caught are indicated by stippling on the main map. Catches include fishes 42.5–180 cm in length and 0.8–95 kg in weight; relative number of landings is shown by different widths of stippling.

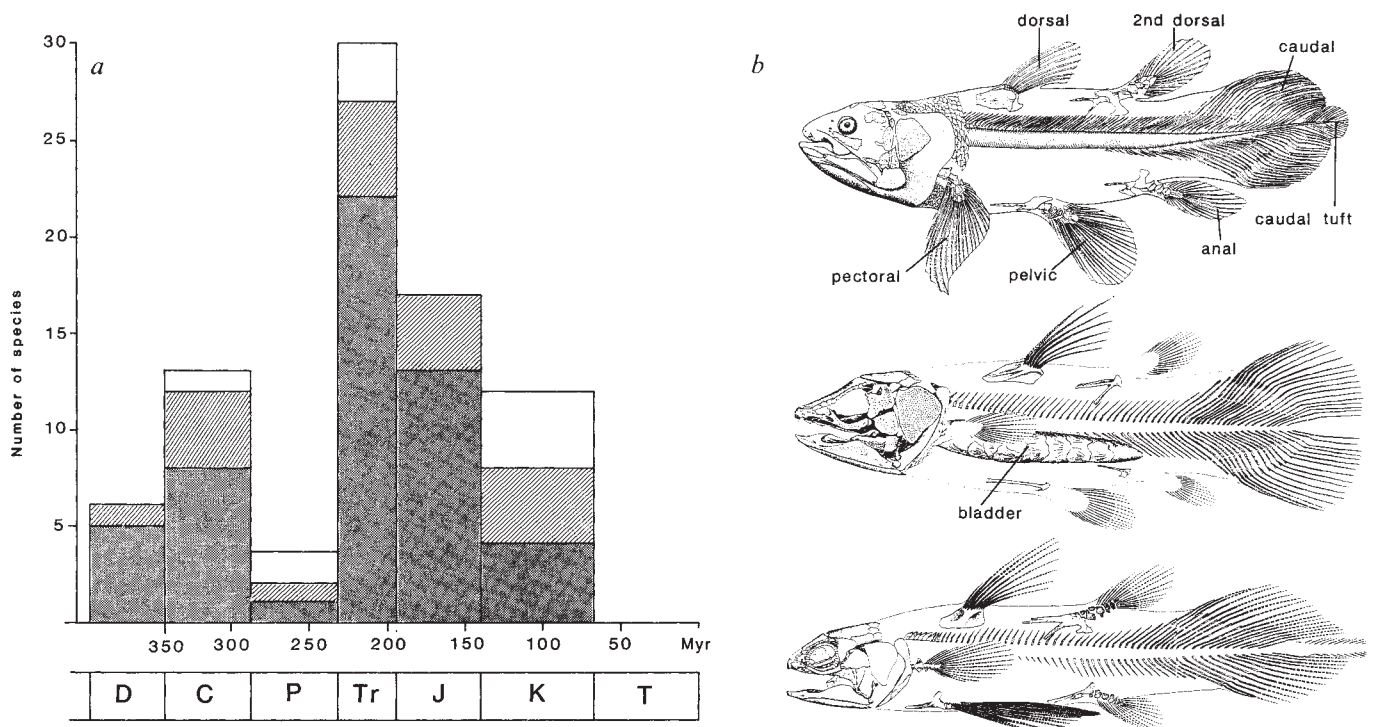


Fig. 2 Diversity of coelacanths through time. *a*, Histogram of numbers of species for each geological period plotted according to environment. Stipple, marine; hatch, freshwater; blank, variable or unknown. D, Devonian; C, Carboniferous; P, Permian; Tr, Triassic; J, Jurassic; K, Cretaceous; T, Tertiary. *b*, Three coelacanths showing their distinctive yet conservative morphology, increasing lobation of the second dorsal and anal fins, and the ossified air bladder of the fossil forms. Top, *Latimeria*; middle, *Macropoma mantelli*, Upper Cretaceous (~80 Myr); bottom, *Laugia groenlandica*, Lower Triassic (~230 Myr).

Darwin, who suggested that living fossils may remain morphologically uniform because they are highly adapted to particular environments and lack competitors; other authors^{21,24} associate slow morphological change with broad ecological adaptation. Coelacanths fit neither model. Darwin's view is supported by the observation that *Latimeria* seems confined to a particular habitat. Apart from the very first specimen, it is

known only from steep and rather barren submarine slopes^{25,26}. On the other hand, coelacanths as a group are known to have lived in a broad range of habitats from coal swamps (*Rhabdoderma*), shallow tropical lagoons (*Undina* from the Jurassic Solnhofen limestones of Bavaria), clear chalk seas (*Macropoma*) to deep oceanic waters (*Latimeria*). Another approach to the problem of slow evolutionary change might be through genetics. But to apply ideas that evolution might be related to generation time²⁷ or differential lineage-specific mutation rate^{28,29} to coelacanths, we need to know a great deal more about the genome of *Latimeria* than we do at present^{20,30,31}.

A test for palaeontological reconstruction

Coelacanths were known as fossils for almost a century before *Latimeria* was discovered. Without an immediate living relative as a guide, skeletal reconstructions were modelled primarily on extinct Devonian fishes termed rhipidistians (see Box) as well as Recent lungfishes and ray-finned fishes. *Latimeria* provided palaeontologists with a yardstick for the reconstructions of coelacanth skeletons based on models from other groups. In most cases these were accurate, although some details of the braincase of *Latimeria*, such as zygial plates⁶, had not been predicted in fossil coelacanths. Furthermore, the reconstruction of fossil coelacanths with rhipidistian features such as internal nostrils (choanae) and maxillae³² was misleading, because *Latimeria* has neither. Many fossil coelacanths have openings immediately in front of the eye, which were interpreted by comparison with ray-finned fishes as posterior external nostrils. But in *Latimeria* these are the posterior openings of a specialized rostral organ, unknown in other fishes.

A missing link between fish and tetrapods

The traditional opinion that *Latimeria* is the closest living relative of tetrapods stems from a palaeontological argument which accepts that both coelacanths and tetrapods are descended from Devonian rhipidistian fishes. This idea had its origins in the

Fishes and their relatives

All the groups listed below belong to the Gnathostomata, or vertebrates with jaws.

Cartilaginous fishes (Chondrichthyes), which have cartilaginous skeletons, include sharks, rays and holocephalans.

Bony fishes (Osteichthyes) include, among others, the groups mentioned below.

- The ray-finned fishes (Actinopterygii) include most extant fishes, the teleosts. *Polypterus* is not a teleost, but a relatively primitive form believed to be related to one of the oldest actinopterygian lineages.
- Sarcopterygii includes the lungfishes; the coelacanths (*Latimeria* and its fossil relatives); and the rhipidistians, an artificial group comprising several kinds of primitive lobe-finned fish, all of which are extinct. They were elongate, with powerful jaws equipped with sharp teeth, and paired pectoral and pelvic fins which had well-developed internal skeletons and were presumably very muscular. They lived from the Devonian to the Permian. Examples include *Osteolepis* and *Eusthenopteron*. Coelacanths and rhipidistians have been grouped together as Crossopterygii. The Tetrapods form a group within the Sarcopterygii that includes all land vertebrates.

Table 1 Characters which have been used to link rhipidistian fishes with tetrapods

1. Intracranial joint immobilised	(L, -C)
2. Choana	(L, -C)
3. Mandible with similar bone patterns	(-C)
4. Septomaxilla	(-C)?
5. Similar pattern of cheek bones	(-C)
6. Similar pattern of skull roofing bones	(-C)
7. Hyomandibular double headed	(L)
8. Tooth enamel folded	(some A)
9. Dorsal ribs	(A, L, -C)
10. Humerus of complex shape, with processes	(L)
11. Screw-shaped glenoid	(L, -C)
12. Vertebrae similarly complex	(Some A)
13. Distribution of dentition	(-C)

The coelacanths were once included with various extinct groups of lobe-finned fishes, collectively termed rhipidistians, in a larger group, the Crossopterygii, which excluded ray-finned fishes (Actinopterygii and lungfishes). Rhipidistians—and by implication coelacanths—were thought to be close to tetrapod ancestry. The characters used to support this theory are given here, but their distribution weakens the argument for a coelacanth-tetrapod affinity as some of them are absent in coelacanths (-C), or present in Actinopterygii (A) or lungfishes (L). (After ref. 35).

1860s when fossil coelacanths were placed with rhipidistians, lungfishes and the primitive African freshwater fish *Polypterus* in the group Crossopterygii on the basis that all these forms possess lobed paired fins³³. This grouping was considered the most closely related to tetrapods because some of them shared similarities in the braincase and others in tooth structure. Towards the end of the nineteenth century, the emphasis shifted from identifying points of similarity towards searching for actual relationships between ancestors and descendants. Then, as now, ancestors were recognized on the basis of primitive characters and rhipidistians were singled out as candidates for tetrapod ancestors. Lungfishes were thought too specialized and *Polypterus* was shown to be related to the Actinopterygii (ray-finned fishes); both were removed from the Crossopterygii.

Against this theoretical backdrop, it was hoped that some trace of rhipidistian (and therefore tetrapod) ancestry would have been frozen in *Latimeria*. Detailed anatomical and functional studies of the fins and the respiratory, circulatory and reproductive systems were expected to elucidate details of the transition of life from water to land. But in many ways, *Latimeria* has proved disappointing as a missing link, or more correctly as an intercalary type—an animal exemplifying primitive and intermediate stages, for it has neither internal nostrils, maxillae nor functional lungs, all to be expected in a rhipidistian derivative and tetrapod relative. The seemingly primitive structure of the brain, heart and intestine are not at all those expected in a tetrapod ancestor. Indeed, some characteristics, such as urea retention and the possession of large eggs, are more like those of cartilaginous fishes (sharks, rays and allies). If *Latimeria* is supposed to be the closest living relative of tetrapods, why are there so few tetrapod-like features?

Palaeontological problems

The evidence of immediate relationship between tetrapods and Crossopterygii (the revised group, containing coelacanths and rhipidistians, but not lungfishes or *Polypterus*) condenses down to the thirteen points of similarity shown in Table 1. It turns out that such evidence as exists to support this relationship is weak. First, it does not support a close relationship between rhipidistians and tetrapods. Several characters have a wider distribution than expected if they are indeed evidence of relationship between the two groups. On reflection this is not surprising because ancestors are wholly primitive with respect to their presumed descendants and primitive characters define wider groups. For instance, the suggestion³⁴ of a close relation-

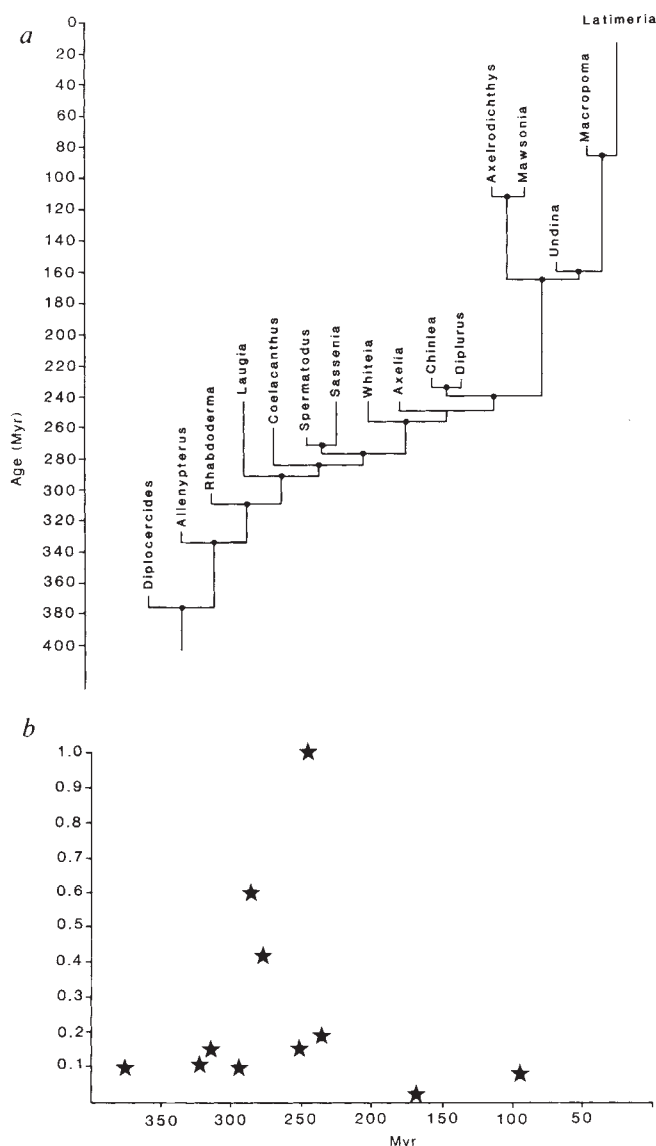
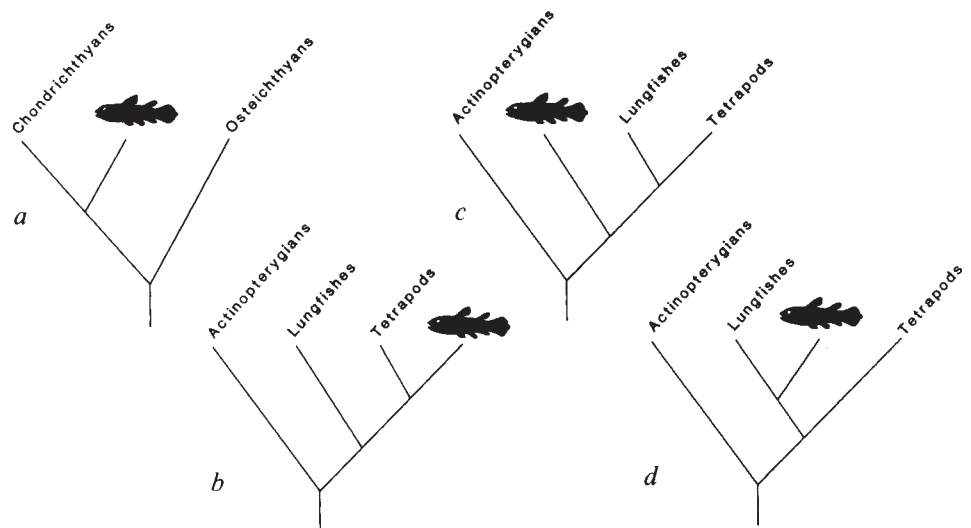


Fig. 3 *a*, Tree of better-known coelacanth genera, based on a cladogram produced from 30 characters using the program PAUP⁶³. Length = 50 steps; consistency index = 0.62. Original data on file in the Department of Palaeontology, British Museum (Natural History). *b* Rate of morphological evolution is given as numbers of new characters acquired between each node in *a* (black dots) divided by the time separating each node. By convention, the origin of a sister-group pair is taken as 5 Myr before the earliest known occurrence of the older sister. This plot implies that the rate of morphological evolution is correlated with the rate of speciation, which reached a maximum in the Triassic (~200 Myr) (Fig. 2).

ship between the primitive Devonian amphibian *Ichthyostega* and certain rhipidistians, which is based on the shared presence of a fish-like tail and opercular bones, is flawed because fish-like tails are characteristic of the general group of vertebrates with jaws (Gnathostomata) whereas opercular bones are characters of bony fishes (Osteichthyes) (see Box). As *Ichthyostega* and rhipidistians are members of both groups these particular characters cannot be used to identify any one group of fishes as the closest relatives of tetrapods.

Second, many of these characters are absent in all coelacanths, which would not be expected were coelacanths derived from rhipidistians. They are usually explained away as peculiarities of the coelacanth lineage. Third, many of the characters are also

Fig. 4 Four theories of the relationships of *Latimeria*. In *a* (refs 40–42), *Latimeria* is considered related to cartilaginous fishes (Chondrichthyes) while in *b* (refs 36, 64), *c* (ref. 35) and *d* (refs 39, 50) it is considered as more closely resembling one or other of the various groups of bony fishes (Osteichthyes), more specifically other representatives of the group that contains lobe-finned fishes and tetrapods (collectively the Sarcopterygii). These theories have all been proposed within a cladistic framework, meaning that attempts to find sister-group relationships take precedence over attempts to link ancestors with descendants. The particular groupings are suggested on the basis of shared derived characters placed at the relevant nodes. These characters are all shared by groups distal to the nodes, and are used to define these groups. Similar characters occurring in groups outside those defined by a particular character represent instances of convergent evolution. The purpose of drawing a cladogram is to find the topology in which the number of characters shared by common descent (homologies) is maximized, and the number of convergences is minimized.



present in lungfishes, which are not included in the Crossopterygii. This implies that any theory that invokes coelacanth and rhipidistians as closest relatives of tetrapods must also include lungfishes in some way. Thus, even if we accept that rhipidistians are closely related to tetrapods, there is no good reason to believe that coelacanth is especially closely related to tetrapods or even rhipidistians.

Latimeria to the rescue

Study of *Latimeria* can augment the palaeontological approach by supplying a wealth of information about the soft parts not preserved in its fossil relatives and that can be compared with features in other extant forms. This information has provoked several theories of relationships between *Latimeria* and other lower vertebrates. One of these proposes that *Latimeria* is most closely related to cartilaginous fishes (Chondrichthyes: the sharks, rays and allies) and so is not very informative about the origins of tetrapods. Three others (Fig. 4) propose that *Latimeria* is a bony fish. As presented here, these theories are condensed from much recent discussion^{35,36–42}.

Latimeria and cartilaginous fishes

This is potentially the most interesting theory (Fig. 4a) because it highlights many of the unexpected features discovered in the living coelacanth. Løvtrup⁴⁰ listed ten soft anatomical and physiological characters common to *Latimeria* and cartilaginous fishes, and Lagios^{41,42} has added characters of the pituitary. Similarities such as large eggs, fatty liver, compact pancreas and the simple structure of the thymus and thyroid are widespread among the lower vertebrates during some or all ontogenetic stages and may be primitive vertebrate features, not significant in suggesting relationship between *Latimeria* and cartilaginous fishes. But similarities in the structure of the pituitary gland^{42,43} are clearly shared derived characters. Whether they are homologous, indicating an immediate relationship, depends on congruence with other characters.

Similarities in the pattern of osmoregulation^{10,41} may also be significant. *Latimeria* and cartilaginous fishes are primarily marine and their body fluids are approximately isotonic with sea water. But the similarities of retention of high levels of urea and trimethylamine oxide in the blood and urine and active excretion of excess salts through a rectal gland may merely be primitive features retained for a marine life. *Latimeria*, however,

shares some features of osmoregulation with bony fishes and tetrapods; in particular, it seems not to resorb urea in the kidneys (in contrast to cartilaginous fishes) and it has a renin–angiotensin system⁴⁴ which in tetrapods is known to be concerned with regulating glomerular filtration rate. The concentration of electrolytes such as Na⁺ and Cl[−] in the blood of *Latimeria* is also very similar to that of marine teleosts¹⁰, but teleosts excrete excess salts across the gills, whereas *Latimeria* uses a rectal gland which resembles most closely that of sharks (a subgroup of the Chondrichthyes). But there are complications; the African lungfish *Protopterus* has a salt-excreting cloacal gland⁴⁵, whereas some chondrichthyans have a gland quite different in structure from those of both *Latimeria* and sharks. The presence of such a gland could be a primitive vertebrate character but, on details of rectal gland anatomy alone, *Latimeria* seems to be a chondrichthyan, and most closely related to sharks. This reasoning implies that *Latimeria* is most closely related to one subgroup of cartilaginous fishes. But theories of relationship which accept that coelacanth is a bony fish are supported by many more characters, and the similarities between *Latimeria* and chondrichthyans, or one subgroup of chondrichthyans, could merely be convergent.

Latimeria and bony fishes

But if coelacanth is related to bony fishes as evidenced by many features of a bony dermal skeleton, what is their position within the group? Two theories^{46,47} suggest that they are more primitive than actinopterygians and only very distantly related to tetrapods. Neither theory is particularly well supported and they are not included in Fig. 4 or discussed here. The remaining theories (Fig. 4b–d) place coelacanth, lungfishes and tetrapods together in a group termed Sarcopterygii. Characters shared by the Sarcopterygii and not known to occur in either actinopterygians or chondrichthyans include limb structure, tooth enamel, details of the gill arches and the presence of an inferior vena cava and pulmonary vein. But it is not certain which of the other sarcopterygian groups is most closely related to *Latimeria*: suggestions have included tetrapods³⁶, lungfishes³⁹ and a combined group of both³⁵.

Latimeria and tetrapods

The relationship to tetrapods (Fig. 4b) is the modern version of traditional ideas, based on fossil representatives of Recent

groups and on exclusively fossil rhipidistians which, according to this theory, are placed on the branch leading to tetrapods. There is surprisingly little evidence for this grouping of *Latimeria* and tetrapods (see above), but there are similarities in the inner ear and enlargements of the spinal cord at the levels of the limbs, both of which deserve further investigation. In the inner ear, both *Latimeria* and tetrapods have a basilar papilla and a perilymphatic duct connecting left and right saccular chambers, although the similarity in detail is greater between the higher vertebrates (amniotes) and *Latimeria* than between either and amphibians. The enlargements of the spinal cord opposite the pectoral and pelvic limbs may have some implication for the coordination of tetrapod locomotion but this feature has not yet been investigated in lungfishes which also make tetrapod-like limb movements.

Latimeria against lungfishes and tetrapods

In the third theory (Fig. 4c), lungfishes are more closely related to tetrapods than *Latimeria*, which is in turn more closely related than any other animal to the joint lungfish-tetrapod group^{35,37}. This is a revitalization of nineteenth-century work which stressed the great similarity between lungfishes and amphibians, the most obvious point of skeletal similarity being the fusion of the palate to the braincase. There are also many similarities in the anatomy of the soft parts, some of which are adaptations to air breathing common to lungfishes and tetrapods⁴⁸ but absent in *Latimeria*.

Some newly described fossils of the primitive lungfish *Diabolichthys* from the Lower Devonian of China⁴⁹ cast doubt on the value of some of the skeletal similarities and weaken support for this theory. Although *Diabolichthys* is clearly a lungfish, the palate was not fused to the braincase and it seems probable that features such as palatal fusion and the position of the internal nostrils are parallelisms between lungfishes and tetrapods. Apart from seriously eroding the lungfish-tetrapod grouping, such parallelisms beg the question of how tetrapods and lungfishes came to share such important structural similarities if it was not through common descent.

Latimeria and lungfishes

The fourth possibility—that *Latimeria* is most closely related to lungfishes (fig. 4d)—has not previously been seriously considered, but support in its favour is accumulating. As proposed by Chang³⁹, various kinds of rhipidistian fishes are included among the close relatives of lungfishes, but in Fig. 4d only the living taxa are shown.

Of the characters shared by *Latimeria* and lungfishes, those associated with the brain⁵⁰ are the most interesting. As in most vertebrates the cerebral hemispheres are evaginated but both *Latimeria* and lungfishes have a prominent septum ependymale dividing left from right. In the diencephalon, the dorsal thalamus, which is concerned with the somatic sensory system in other vertebrates, is very thick. It is of interest to note that both lungfishes and *Latimeria* have a well-developed system of electroreceptor cells over the snout⁵¹, although a direct functional link between electroreceptive organs and the dorsal thalamus remains to be demonstrated.

The third and fourth theories are, in my view, equally worthy of consideration with little to choose between them on current information. Amino-acid and nucleotide sequence data are possible avenues for investigation in the immediate future, although the evidence so far is contradictory. Whereas amino-acid sequence data from α - and β -parvalbumin⁵² show a closer relationship between actinopterygians and tetrapods than between either and *Latimeria*, the reverse appears to be true from analysis of a partial nucleotide sequence of 28s rRNA³¹. Considerably more molecular information on *Latimeria* and other animals is needed if it is to be useful for resolving competing theories of relationship.

Latimeria alone

Neither the third nor the fourth theories support *Latimeria* as the nearest living relative of tetrapods, or as a missing link, and it is certainly over-optimistic to expect that *Latimeria* will yield secrets of the fish-amphibian transition. Indeed, there are remarkably few features which link *Latimeria* with tetrapods; and *Latimeria* is best considered as a bony fish which has some features closely resembling those in chondrichthyans (presumably through convergence) as well as retaining many primitive vertebrate features. But in other respects it seems uniquely specialized, in particular the ventral location of the kidney, the rostral organ and the diminutive brain⁵³. Some features are a result of adaptation to life in deep water at relatively low temperatures, including the rod-dominated retina, the low gill-surface area, the particular oxygen dissociation curve and, possibly, the fat-filled gas bladder. In yet other respects, *Latimeria* may retain juvenile features, which may mask features that we would expect in a tetrapod relative. Such paedomorphic features include the simple thymus and thyroid as well as the seemingly embryonic heart, which lacks the 'S' curvature of most adult vertebrate hearts. Further support for paedomorphism comes from the reduction of ossification, particularly of the endoskeleton, which in phylogenetic terms can come about only by successive changes in the timing of ossification during development.

Special features of *Latimeria* biology

Three features of the biology of *Latimeria* are of particular interest: the intracranial joint, the limbs, and the possible use of electroreception. The intracranial joint, which potentially allows the skull to flex (Fig. 5) was present in many Palaeozoic lobe-finned fishes³⁷, whether or not it was subsequently lost in

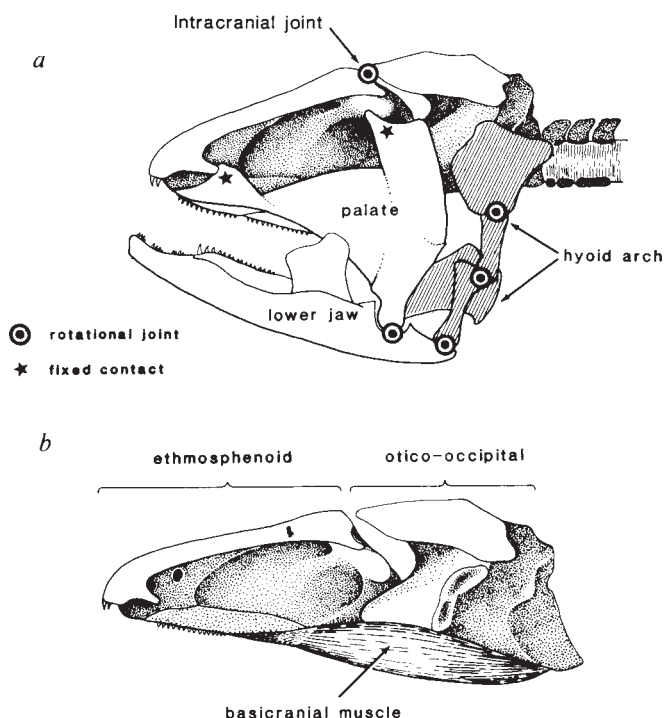


Fig. 5 Intracranial joint of *Latimeria*. *a*, Skull with cheek bones removed to show braincase, mandibular (palate and lower jaw) and hyoid arches. The joint is closed by the action of a powerful basicranial muscle. *b*, Braincase showing position of this muscle. Manipulation of freshly-dead specimens⁶⁵ shows that the ethmosphenoid can flex vertically through about 15° with respect to the otico-occipital. The palate is securely attached to the ethmosphenoid and the two move as a unit, independent of the hyoid arch. Together with the lower jaw, the entire mechanism acts as a four-bar linkage system⁶⁶.

tetrapods. *Latimeria* is the only animal alive today that has such a joint. What it is for and how it works are still matters for debate⁵⁴: in extinct lobe-finned fishes it may have acted as a shock absorber during prey capture; increased the gape; or even allowed expansion of the buccal cavity when the mouth was closed. Examination of *Latimeria* by filming live animals and making detailed electromyographic recordings⁵⁵ to follow the sequence of contraction of the different muscles concerned with the joint could decide the issue. Through geological time, the front half of the coelacanth skull (ethmosphenoid) has become relatively longer than the back half (otico-occipital), potentially increasing the gape and the relative length of the basicranial muscle that closes the joint. The significance of these changes may become apparent once the mechanism is understood in *Latimeria*.

The structure and function of the fins are potentially informative for our understanding of tetrapod locomotion, and their movements have recently been documented²⁶. Both pectoral and pelvic fins are very mobile and show a great deal of rotational movement. The pectoral fin can rotate through 180° on its axis, as predicted for a precursor of the tetrapod limb. Of greater interest is that, during slow forward movements, the synchronicity of abduction and adduction is decidedly tetrapod-like, with the left pectoral and right pelvic fins moving in phase and opposing the right pectoral and left pelvic fins²⁶. But predictions⁴ that the coelacanth can walk on the substrate have not been confirmed. The details of the neuromuscular control of the paired fins will be of great interest, especially if comparisons can be made with lungfishes which have similar limb coordination³⁵. Interestingly, the unpaired second dorsal and anal fins (Fig. 2b) are supported by an endoskeleton very similar to that of the paired fins; they are very mobile and exhibit synchronous lateral and rotational sculling movements.

Electroreception with ampullary organs on the head is probably a primitive vertebrate feature⁵⁶, although its elaboration represents a derived condition. Lungfishes are probably specialized as they have electroreceptors on the trunk as well as the head⁵⁶ and it is very likely that *Latimeria*, which lives in darkness, is also highly specialized with electroreceptors concentrated in the

rostral organ on the snout⁵¹. Observations of *Latimeria* from a submersible²⁵ show that it performs a peculiar head-standing behaviour that can be induced by creating a weak electric field⁵⁷.

Conservation

Just how successfully we will be able to increase our understanding of these functions will, in part, depend upon the conservation of *Latimeria*, which gives cause for concern. Since 1938 there have been approximately 200 recorded catches¹⁴, most by Comoran fisherman using long lines and operating from small dugout canoes. More sophisticated craft are increasingly being used⁵⁸, able to operate in deeper waters and for longer periods, and there is evidence that the catch has increased dramatically in the past few years above the previously steady three or four a year. We simply do not know how long this increase can be sustained before irreparable damage is done to the coelacanth population. We know nothing of the population size, age or sex structure (recorded sex of catches is approximately equal male and female^{59,60}), longevity, age at maturity or fecundity. The fish is ovoviparous⁶¹ and fecundity is probably low as, in the single pregnant female known, only five advanced fetuses were found occupying almost the entire oviduct; it is likely that gestation time exceeds 12 months⁶². Most recorded catches have been of medium and large individuals (longer than 1 m) and we know practically nothing of the young, where they live, their diet⁶⁰ or other habitat requirements. Nor do we yet know their geographic range, although the apparent preference for cold water and steep submarine slopes may limit their distribution^{14,25}. There have been anecdotal reports of coelacanth catches off the coast of East Africa, and also in the Mediterranean, but so far without confirmation.

The primary concern of the newly founded Coelacanth Conservation Council^{14,58} is to correct these significant deficiencies in our knowledge, and we can only hope that the celebration of the Golden Jubilee heralds a new era for the appreciation of this unique steel-blue fish.

I thank Drs C. Patterson and P. H. Greenwood, British Museum (Natural History) for reading drafts of this article and for their helpful comments.

1. Courtenay-Latimer, M. *Occ. Pap. Calif. Acad. Sci.* **134**, 6-10 (1979).
2. Smith, J. L. B. *Nature* **143**, 455-456 (1939).
3. Smith, J. L. B. *Nature* **171**, 99-101 (1953).
4. Smith, J. L. B. *Old Fourlegs. The Story of the Coelacanth* (Longman Green, London, 1956).
5. Smith, J. L. B. *Trans. R. Soc. S. Afr.* **28**, 1-106 (1940).
6. Millot, J. & Anthony, J. *Anatomie de Latimeria chalumnae I: Squelette et Muscles*. (CNRS, Paris, 1958).
7. Millot, J. & Anthony, J. *Anatomie de Latimeria chalumnae II: Système Nerveux et Organes des Sens*. (CNRS, Paris, 1965).
8. Millot, J., Anthony, J. & Robineau, D. *Anatomie de Latimeria chalumnae III* (CNRS, Paris, 1978).
9. Northcutt, R. G., Neary, T. J. & Senn, D. G. *J. Morph.* **155**, 181-192 (1978).
10. Griffith, R. W. & Pang, P. K. T. *Occ. Pap. Calif. Acad. Sci.* **134**, 79-93 (1979).
11. Hughes, G. M. *Proc. R. Soc. Lond. B208*, 309-328 (1980).
12. Griffith, R. W. *Proc. R. Soc. Lond. B208*, 329-347 (1980).
13. Locket, N. A. *Proc. R. Soc. Lond. B208*, 265-307 (1980).
14. Balon, E. K., Bruton, M. N. & Fricke, H. *Environ. Biol. Fishes* **23**, 241-279 (1988).
15. Stanley, S. M. *Macroevolution: Pattern and Process* (Freeman, San Francisco, 1979).
16. Eldredge, N. E. & Stanley, S. M. in *Living Fossils* (eds Eldredge, N. E. & Stanley, S. M.) 1-3 (Springer, New York, 1984).
17. Schaeffer, B. *Bull. Am. Mus. nat. Hist.* **99**, 25-78 (1952).
18. Westoll, T. S. in *Genetics, Paleontology and Evolution* (eds Jepsen, G. L., Simpson, G. G. & Mayr, E.) 121-184 (Princeton University Press, Princeton, 1949).
19. Simpson, G. G. *Tempo and Mode in Evolution* (Columbia University Press, New York, 1944).
20. Goodman, M., Miyamoto, M. M. & Czelusniak, J. in *Molecules and Morphology in Evolution: Conflict or Compromise?* (ed. Patterson, C.) 141-176 (Cambridge University Press, Cambridge, 1987).
21. Eldredge, N. E. *Bull. Carnegie Mus. nat. Hist.* **13**, 7-19 (1979).
22. Schaeffer, B. *Am. Mus. Novit.* **1110**, 1-17 (1941).
23. Forey, P. L. *Palaeontology* **24**, 203-229 (1981).
24. Simpson, G. G. *The Major Features of Evolution* (Columbia University Press, New York, 1953).
25. Fricke, H. & Plante, R. *Naturwissenschaften* **75**, 149-151 (1988).
26. Fricke, H., Reinicke, O., Hofer, H. & Nachtigall, W. *Nature*, **329**, 331-333 (1987).
27. Sibley, C. G. & Ahlquist, J. E. in *Molecules and Morphology in Evolution: Conflict or Compromise?* (ed. Patterson, C.) 95-121 (Cambridge University Press, Cambridge, 1987).
28. Dobzhansky, T. *Genetics and the Origin of Species* 2nd edn (Columbia University Press, New York, 1941).
29. Britten, R. *Science* **231**, 1393-1398 (1986).
30. Dingerkus, G. *Occ. Pap. Calif. Acad. Sci.* **134**, 111-127 (1979).
31. Hillis, D. in *The Hierarchy of Life*. (Excerpta Medical International Congress Series, Elsevier, Amsterdam, in the press).
32. Nielsen, E. *Meddr. Grønland* **112**, 1-55 (1936).
33. Huxley, T. H. *Mem. geol. Surv. U.K.* **10**, 1-40 (1861).
34. Jarvik, E. *Meddr. Grønland* **114**, 1-90 (1952).
35. Rosen, D. E., Forey, P. L., Gardiner, B. G. & Patterson, C. *Bull. Am. Mus. nat. Hist.* **167**, 159-276 (1981).
36. Schultze, H.-P. *J. Morph. Suppl.* **1**, 39-74 (1987).
37. Forey, P. L. *J. Morph. Suppl.* **1**, 75-91 (1987).
38. Panchen, A. L. & Smithson, T. R. *Biol. Rev.* **62**, 341-438 (1987).
39. Chang, M. M. in *Origins of the Higher Categories of Tetrapods* (ed. Trueb, L., Schultze, H.-P. & Martin, L.) (Museum of Natural History, Kansas University, Lawrence, in the press).
40. Löwtrup, S. *The Phylogeny of the Vertebrata* (Wiley, London, 1977).
41. Lagios, M. D. *Occ. Pap. Calif. Acad. Sci.* **134**, 25-44 (1979).
42. Lagios, M. D. *Copeia* **1982**, 942-948 (1982).
43. Robineau, D. *Ann. Sci. nat. Zool.* **13** ser. **8**, 43-60 (1987).
44. Lagios, M. D. *Gen. comp. Endocr.* **22**, 296-307 (1974).
45. Lagios, M. D. & McCosker, J. E. *Copeia* **1977**, 176-178 (1977).
46. Wiley, E. O. *Zool. J. Linn. Soc.* **67**, 149-179 (1979).
47. Wahlert, G. von *Latimeria und die Geschichte der Wirbeltiere* (Fischer, Stuttgart, 1968).
48. Burgren, W. W. & Johansen, K. J. *Morph. Suppl.* **1**, 217-236 (1987).
49. Chang, M. M. & Yu, X. *Proc. Linn. Soc. S.W.* **107**, 171-184 (1984).
50. Northcutt, R. G. *J. Morph. Suppl.* **1**, 277-297 (1987).
51. Bemis, W. E. & Hetherington, T. E. *Copeia* **1982**, 467-471 (1982).
52. Maeda, N., Zhu, D. & Fitch, W. M. *Molec. Biol. Evol.* **1**, 473-488 (1984).
53. Anthony, J. & Robineau, D. C. r. *hebd. Séanc. Acad. Sci., Paris* ser. D. **283**, 1739-1742 (1976).
54. Thompson, K. S. *Zool. J. Linn. Soc.*, **46**, 223-253 (1967).
55. Lauder, G. V. *J. Morph.* **163**, 283-317 (1980).
56. Northcutt, R. G. in *Electroreception* (ed. Bullock, T. H. & Heiligenberg, W.) 257-285 (Wiley & Sons, New York, 1986).
57. Fricke, H. *Natn. geogr. Mag.* **173**, 824-838 (1988).
58. Bruton, M. *Ichthos* **18**, 1-2 (1988).
59. Anthony, J. *Operation Coelacanth* (Arthaud, Paris, 1976).
60. McCosker, J. E. *Occ. Pap. Calif. Acad. Sci.* **134**, 17-24 (1979).
61. Smith, C. L., Rand, C. S., Schaeffer, B. & Atz, J. W. *Science* **190**, 1105-1106 (1975).
62. Hureau, J. C. & Ozouf, C. *Cybum ser.* **3**, 129-137 (1977).
63. Swofford, D. *Phylogenetic Analysis Using Parsimony. User's Manual* (Illinois Natural History Survey, Champaign, Illinois, 1985).
64. Fritsch, B. *Nature* **327**, 153-154 (1987).
65. Thompson, K. S. *Postilla* **149**, 1-12 (1970).
66. Lauder, G. V. *Copeia* **1980**, 1-9 (1980).

Decoupled evolution of Nd and Sr isotopes in the continental crust and the mantle

Steven L. Goldstein

Max-Planck-Institut für Chemie, Saarstrasse 23, Postfach 3060, D-6500 Mainz, FRG

Nd and Sr isotopes evolve independently in clastic sedimentary rocks and the continental crust as a whole, because Sr is recycled preferentially from the continents to the mantle. As a result, the differentiation age of the mantle inferred from Sr isotopes and the mean Rb–Sr age of the continental crust are younger than their true ages.

EVALUATION of the overall Rb–Sr isotope evolution of the continental crust is complicated because Rb/Sr ratios are easily fractionated by common crustal processes such as weathering, melting and metamorphism, and because Sr resides in carbonates as well as in silicates. In contrast, Sm and Nd are resistant to disturbance by these processes, and reside almost totally in silicates. Because of their simpler geochemical behaviour, Sm–Nd isotope studies have been the preferred means to investigate the history of continental growth and mantle differentiation.

Here I show that Sm–Nd studies, particularly of clastic sediments, provide a basis for understanding Rb–Sr isotopes in the continental crust, and that the age of the continental crust implied by Rb–Sr studies is young compared with that derived from Sm–Nd data. Recognition of this relationship and understanding of its causes allows a new interpretation of Rb–Sr isotope data on continental rocks, and has important implications for the evolution of Sr isotopes in the mantle.

Nd isotopes in sediments

To understand Nd isotopes in clastic sediments it must be realized that half of existing sediments have been deposited within the past 250–500 Myr (refs 1–3). The young depositional age of sediments contrasts with the continental crust, whose mean age is probably >2 Gyr and whose rate of growth today is lower than in the past⁴. The (depositional) youth reflects loss of old sediments, which are destroyed through erosion and re-deposition^{1–3}. For example, Veizer and Jansen^{2,3} estimate that up to 95% of the material in newly deposited sediments comes from eroded old sediments.

Nd isotopes in Phanerozoic sediments are consistent with this interpretation. Figure 1a shows that the present-day ϵ_{Nd} of Palaeozoic and Mesozoic shales (averaging -11 to -13) are similar to Recent sediments derived from large areas of the continental crust (for example, from major rivers, wind-borne dusts and loess)^{5,6}. From a consideration of only the Nd isotopes, the Recent sediments could be formed by erosion and re-deposition of older sediments with the possibility of a small additional input of young continental material (constituting for example, $\sim 10\%$ if $\epsilon_{\text{Nd}} \approx +8$). If these sediments contain significant components from non-sedimentary sources, such as old or young crystalline rock with ϵ_{Nd} very different from -12 , then they must be well mixed before deposition and have average $\epsilon_{\text{Nd}} \approx -12$.

Although the average depositional age of the sedimentary mass is young, the material in the sediments has been in the continental crust much longer, ~ 1.7 Gyr on average for the Nd in present-day wind-blown dusts and major-river sediments⁶. If half of the sedimentary mass has been deposited within the past 250–500 Myr, the material in a Recent sediment could have experienced as many as 3–7 erosion–deposition cycles. The difference between depositional ages and ‘crustal residence’ ages⁷, combined with the uniformity of ϵ_{Nd} , provides further evidence for a nearly closed, well mixed sedimentary system

dominated by cannibalistic recycling (erosion and re-deposition).

Sr isotopes in sediments

Figure 1b shows that major river sediments and loess mostly have present-day $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of 0.710–0.720, whereas few Palaeozoic shales have ratios <0.72 . The Sr in Recent sediments cannot be derived simply from erosion of old sediments, and at first glance the Sr isotope data appear to preclude interpretation of the sedimentary mass as a nearly closed system dominated by rapid internal recycling.

Compared with most rock types, shales have high Rb/Sr ratios^{8–12}, typically >1 . If the sedimentary mass were a closed system with respect to Rb–Sr, the high Rb/Sr would result in a marked increase in the initial $^{87}\text{Sr}/^{86}\text{Sr}$ with time: if Rb/Sr >1 and $^{87}\text{Sr}/^{86}\text{Sr} \approx 0.710$, then within 500 Myr $^{87}\text{Sr}/^{86}\text{Sr} > 0.730$ –0.740. Figure 1c shows that there has been no secular increase in initial $^{87}\text{Sr}/^{86}\text{Sr}$ in Phanerozoic shales, but instead the initial $^{87}\text{Sr}/^{86}\text{Sr}$ has remained remarkably constant. The contrast between the constancy of initial values and the huge variations in the present day (Fig. 1b) demonstrates that the present-day values are a result of post-depositional *in situ* decay of ^{87}Rb . The constancy of initial $^{87}\text{Sr}/^{86}\text{Sr}$, despite high Rb/Sr, means that the Sr isotope compositions are buffered. Whereas Nd isotopes might be buffered by Nd in old sediments, the Sr isotopes must be buffered from outside the clastic sediment system.

Figure 1 contains only Phanerozoic data because determination of initial $^{87}\text{Sr}/^{86}\text{Sr}$ in Precambrian sediments is complicated by post-depositional fractionation of Rb/Sr. Apparent initial $^{87}\text{Sr}/^{86}\text{Sr}$ values within a single data set^{13–15} often include both high values and values lower than those in the initial Earth. Because the Phanerozoic accounts for a major portion (50–80%) of the sedimentary mass^{1–3}, the data in Fig. 1 adequately characterize Sr isotopes in clastic sediments. In any case, studies of Precambrian clastic sediments^{13,14} indicate that initial $^{87}\text{Sr}/^{86}\text{Sr}$ values have averaged ~ 0.71 –0.72, despite high Rb/Sr since at least the middle Proterozoic. Figure 1b and c shows that 0.710–0.720 is a small range in the context of $^{87}\text{Sr}/^{86}\text{Sr}$ in shales. This figure includes sediment data from all over the Earth, demonstrating that the constancy of initial $^{87}\text{Sr}/^{86}\text{Sr}$ and the high rate of post-depositional radiogenic growth are global phenomena.

Processes controlling Sr in sediments

The closed-system behaviour of Nd and open behaviour of Sr reflect their responses to weathering. Sr is leached from silicates by water, whereas Nd and Sm are insoluble. Rb is soluble but has a high affinity for clays¹⁶. As a result clastic sediments have high Rb/Sr, low Sr concentrations, and Sr/Nd typically ~ 4 (refs 8–12). In contrast, orogenic basalts and andesites have much higher Sr/Nd of ~ 40 (refs 11, 17). Figure 2 shows that mechanical addition of as little as 5–10% of young volcanic

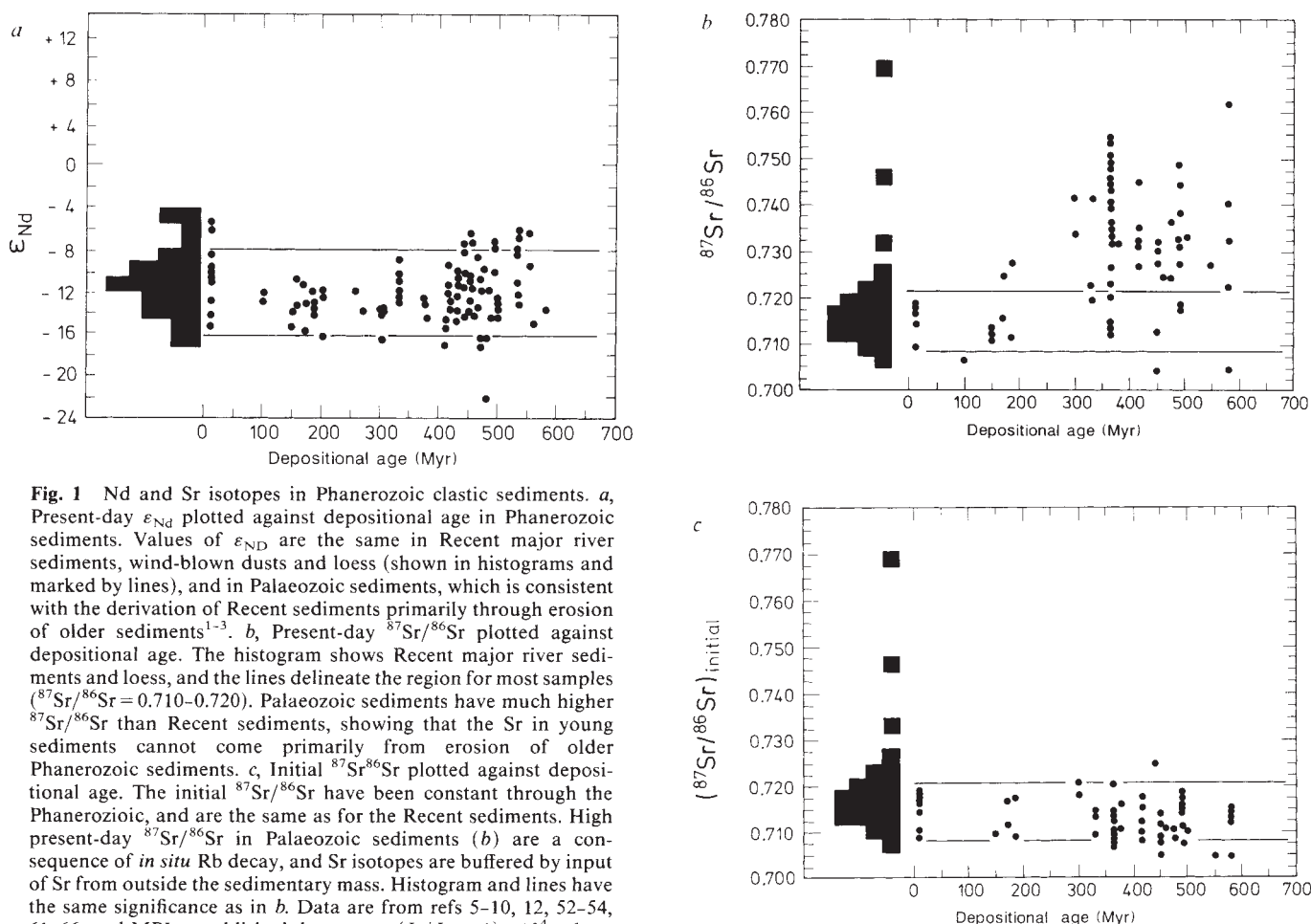


Fig. 1 Nd and Sr isotopes in Phanerozoic clastic sediments. *a*, Present-day ϵ_{Nd} plotted against depositional age in Phanerozoic sediments. Values of ϵ_{Nd} are the same in Recent major river sediments, wind-blown dusts and loess (shown in histograms and marked by lines), and in Palaeozoic sediments, which is consistent with the derivation of Recent sediments primarily through erosion of older sediments¹⁻³. *b*, Present-day $^{87}Sr/^{86}Sr$ plotted against depositional age. The histogram shows Recent major river sediments and loess, and the lines delineate the region for most samples ($^{87}Sr/^{86}Sr = 0.710-0.720$). Palaeozoic sediments have much higher $^{87}Sr/^{86}Sr$ than Recent sediments, showing that the Sr in young sediments cannot come primarily from erosion of older Phanerozoic sediments. *c*, Initial $^{87}Sr/^{86}Sr$ plotted against depositional age. The initial $^{87}Sr/^{86}Sr$ have been constant through the Phanerozoic, and are the same as for the Recent sediments. High present-day $^{87}Sr/^{86}Sr$ in Palaeozoic sediments (*b*) are a consequence of *in situ* Rb decay, and Sr isotopes are buffered by input of Sr from outside the sedimentary mass. Histogram and lines have the same significance as in *b*. Data are from refs 5-10, 12, 52-54, 61-66, and MPI unpublished data. $\epsilon_{Nd} = (I_a/I_{BE} - 1) \times 10^4$, where I is the $^{143}Nd/^{144}Nd$, a refers to the sample and BE to the bulk Earth. $I_{BE} = 0.512638$ (ref. 67).

detritus can bring the $^{87}Sr/^{86}Sr$ value of a shale from >0.730 to <0.720 , whereas the $^{143}Nd/^{144}Nd$ value is hardly affected. What actually occurs within a river is more complicated, because weathering progresses during transport. It is clear, however, that a small amount of orogenic volcanic debris added to a shale has a great effect on Sr isotopes without significantly affecting Nd.

Figure 2 illustrates a situation in which the Nd isotopes are buffered by internal recycling whereas the Sr isotopes are buffered by an input from orogenic volcanics. Both silicates and carbonates contribute to the load of dissolved Sr, and carbonates¹⁸ generally have $^{87}Sr/^{86}Sr < 0.710$. Exchange between dissolved and particulate Sr lowers the $^{87}Sr/^{86}Sr$ in the particulates and thereby enhances the buffering of the Sr isotopes. Although the amount of exchange that occurs is not well constrained, it appears to be insignificant in sea water^{8,9}, with relatively high concentrations of dissolved Sr, and is therefore unlikely to be significant in rivers. The low concentrations of Nd in waters ensure that exchange affects $^{143}Nd/^{144}Nd$ of the water, and not the particulates.

Sm-Nd and Rb-Sr ages in the continental crust

Sm-Nd crustal residence ages correspond to approximate estimates of average residence times in the continents, and therefore can be used to evaluate Rb-Sr ages in the continental crust. For example, if the Sm-Nd and Rb-Sr ages of the continental crust were the same, a sample of average continent would give the same crustal residence age for both systems. There is no *a priori* reason however, for the ages to be equal. Rb is much less compatible with mantle minerals than is Sr, whereas Sm is only

slightly more compatible than Nd. If mantle melting is the ultimate source of the continents, then Rb is transported to the continental crust faster than Sr, Sm and Nd, and Rb/Sr in the mantle is fractionated earlier than Sm/Nd. Isotope ages reflect the timing of parent-daughter fractionation²⁵. Based solely on melting behaviour it would be expected that the Rb-Sr age of the continents is older than Sm-Nd, and this result is found by mantle-crust evolution models²⁵⁻²⁷. It is shown below that the Rb-Sr age is in fact younger than the Sm-Nd age.

The crustal-residence-age equation used in many studies^{6,12,53,54} is $T_j = (1/\lambda) \ln[1 + (I_r - I_s)/(P_r - P_s)]$, where T refers to the 'age', j denotes the decaying element, I_r and I_s are the isotopic compositions of, respectively, a rock and a sample of the depleted mantle system from which the continental crustal material was ultimately derived, P_r and P_s are the parent-daughter ratios, and λ is the decay constant. Weathering causes Rb/Sr in silicates to increase (and therefore $P_r - P_s$ increases and T_j decreases), but has little effect on Sm/Nd. As a result, Rb-Sr crustal-residence ages (T_{Sr}) generally have no true age significance, and are younger than Nd ages (T_{Nd}). Figure 1b and c shows that the rates of increase of $^{87}Sr/^{86}Sr$ in shales subsequent to deposition are so high that T_{Sr} are only a few hundred million years, much younger than the T_{Nd} of $\sim 1.7-1.9$ Gyr (ref. 6).

The anomalously young T_{Sr} of the sediments cannot be attributed solely to an increase of Rb/Sr by weathering. The crustal-residence-age equation also gives young ages if $I_r - I_s$ is close to zero. If it is assumed that the T_{Nd} of ~ 1.8 Gyr in major-river sediments and Phanerozoic shales imply true crustal residence times of $\sim 1.5-2.0$ Gyr, and if the Rb/Sr of the upper

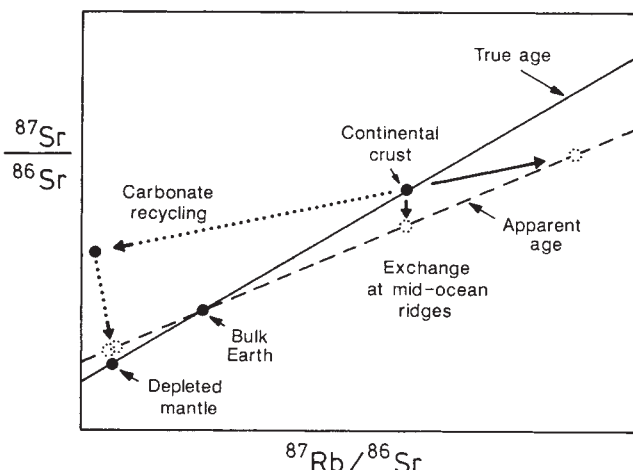


Fig. 4 Rb-Sr age of the mantle-continent differentiation and effects of recycling of ^{87}Sr . If average $^{87}\text{Sr}/^{86}\text{Sr}$ and Rb/Sr were known for the continental crust and its mantle complement, these values and those of the bulk Earth would lie on a line on a Rb-Sr isochron diagram indicating the mean Rb-Sr age of mantle-continent differentiation. Recycling of continental Sr to the mantle by equimolar exchange of seawater Sr with basalt would increase the $^{87}\text{Sr}/^{86}\text{Sr}$ in the mantle and decrease it in the continental crust, leaving Rb/Sr unchanged. A line with a shallower slope results, corresponding to an apparent age that is younger than the true age. If carbonate, with high Sr but low Rb/Sr is recycled to the mantle, then there is a net additional flux of Sr to the mantle. The Rb/Sr and $^{87}\text{Sr}/^{86}\text{Sr}$ of the continental crust increase, resulting in a shallower slope and again a too-young apparent age.

value ~ 0.715 as a result of addition of Sr with $^{87}\text{Sr}/^{86}\text{Sr} \approx 0.705$ from this reservoir, then mass balance requires that 33–70% of the Sr in the upper crust must come from the xenolith reservoir. If $^{87}\text{Sr}/^{86}\text{Sr}$ has been raised from ~ 0.705 (similar to orogenic or continental basaltic magmas, the probable sources of the meta-igneous xenoliths) to ~ 0.707 in the xenolith reservoir by influx of Sr from the upper crust, the amount of upper-crustal Sr in this reservoir is only ~ 6 –15%. Therefore a massive flux of Sr from the lower to the upper crust is required, whereas a significant reverse flux is precluded. This would result in a huge increase with time of the Sr concentration in the upper crust, for which there is no evidence. Although buffering of upper-crustal $^{87}\text{Sr}/^{86}\text{Sr}$ by lower crust cannot be dismissed, only the reservoir represented by the meta-igneous xenoliths is a viable candidate, and an *ad hoc* process is required that allows unidirectional upward migration of Sr within crust. Intra-crustal fractionation appears unlikely to account for the young T_{Sr} and low $^{87}\text{Sr}/^{86}\text{Sr}$ in the upper crust, and it is more likely that the Rb-Sr age of the whole continental crust is too young.

Recycling of ^{87}Sr to the mantle

What processes cause the Rb-Sr age of the continental crust to be younger than the Sm-Nd age? It has been shown that hydrothermal exchange of sea water and basalt at spreading ridges facilitate recycling of elements such as Mg from the continents to the mantle^{28–30}. It is generally accepted that $^{87}\text{Sr}/^{86}\text{Sr}$ in sea water is balanced by inputs from the continents and oceanic basalts^{31,32}. These processes can also significantly affect the Sr isotopic evolution of the continents and mantle. Albarede *et al.*³³ estimated that in seafloor hydrothermal systems there is nearly equimolar exchange of $\sim 10^{10}$ mol yr^{-1} of Sr. Exchange between present-day seawater ($^{87}\text{Sr}/^{86}\text{Sr} = 0.709$) and spreading-ridge basalts ($^{87}\text{Sr}/^{86}\text{Sr} \approx 0.7025$) results in a loss of 0.6×10^7 mol yr^{-1} of ^{87}Sr from sea water. The continent production rate of ^{87}Sr is 3×10^7 mol yr^{-1} if the Rb concentration is ~ 32 p.p.m. (ref. 11) and the mass of the continental crust³⁴ is 2×10^{25} g. If the Sr in the basalts is returned to the mantle, the

annual loss of ^{87}Sr from the continents is $\sim 20\%$ of the production by Rb decay.

In contrast to Sr, recycling of rock to the mantle is the only means for returning significant amounts of Nd and Pb—these elements are nearly insoluble in water, and seawater alteration has negligible effects on Nd and Pb isotopes of mid-ocean-ridge basalt (MORB)³⁵. If hydrothermal exchange is the only process that recycles Sr to the mantle preferentially to Nd, if the present Sr exchange rate of $\sim 10^{10}$ mol yr^{-1} has been typical throughout Earth's history, and if Sm-Nd and Rb-Sr ages otherwise be equal, then the apparent Rb-Sr age in the continental crust is more than 10% lower than the Sm-Nd age. The integrated effect accounts for lower $^{87}\text{Sr}/^{86}\text{Sr}$ in sea water in the past¹⁸. To the extent that other processes also facilitate preferential recycling of continental Sr to the mantle—for example, carbonate subduction^{36,37} or seawater Sr addition to MORB during low-temperature alteration³⁸—then the apparent Rb-Sr age of the continents would be further lowered.

Figure 4 illustrates why recycling of continental Sr to the mantle through ridge hydrothermal processes or carbonate subduction lowers the apparent Rb-Sr age of the continental crust. If the Rb/Sr and $^{87}\text{Sr}/^{86}\text{Sr}$ could be measured for samples of average continent and its mantle complement, they would be colinear with the bulk Earth on an Rb-Sr isochron diagram, and the implied age would be the mean Rb-Sr age of continent-mantle differentiation. Equimolar exchange of seawater Sr at ocean ridges would not effect Rb/Sr, but would increase the $^{87}\text{Sr}/^{86}\text{Sr}$ in the basalt while lowering it in the continent, which decreases the slope of the line and also implies a younger age. Recycling of ocean-floor carbonate increases $^{87}\text{Sr}/^{86}\text{Sr}$ of the mantle and adds Sr. This causes increases of Rb/Sr and $^{87}\text{Sr}/^{86}\text{Sr}$ in the continent because $^{87}\text{Sr}/^{86}\text{Sr}$ in the subducted carbonates are lower than in average continent. The result is again a line with a shallower slope and therefore implies a younger apparent age—the average $^{87}\text{Sr}/^{86}\text{Sr}$ of the continents increases, but is too low compared with the Rb/Sr.

The effects of hydrothermal exchange on the $^{87}\text{Sr}/^{86}\text{Sr}$ evolution of the mantle depend on seafloor spreading rates, $^{87}\text{Sr}/^{86}\text{Sr}$ of sea water, and the rate that Sr has been depleted in the mantle. Figure 5 shows the magnitude of the effect on an upper mantle extending to 670 km ($\sim 28\%$ of the whole mantle³⁹), assuming that Sr has been extracted at a uniform rate since 3.8 Gyr. The amount of ^{87}Sr recycled to the mantle is estimated by assuming an exchange of 10^{10} moles of seawater Sr (refs 18, 40) per 17 km^3 of oceanic crust (the present-day rates^{4,33}). The $^{87}\text{Sr}/^{86}\text{Sr}$ curves of Fig. 5 show only the effects of hydrothermal exchange, that is, as if there has been no other source of radiogenic ^{87}Sr in the upper mantle. Curve 1 is a maximum estimate, for which seafloor spreading rates are proportional to the square of heat production⁴¹ so that basalt production rates were ~ 10 times higher than today at 3.5 Gyr BP. Hydrothermal exchange alone would account for over 3/4 of the increase of $^{87}\text{Sr}/^{86}\text{Sr}$ in the source of MORB. Curve 2 assumes that spreading rates were half of that in curve 1, and exchange would account for over half of the increase. Curve 3 assumes constant spreading rates, and still accounts for $\sim 1/3$ of the increase.

Figure 5 shows that a significant portion of the $^{87}\text{Sr}/^{86}\text{Sr}$ increase in the upper mantle might be a result of Sr recycling from the continents through hydrothermal exchange between sea-water and MORB. The effects would be smaller if exchanged Sr is incorporated into arc magmas or lithospheric mantle, if seafloor spreading rates were lower in the past, or if the depleted portion of the mantle is deeper than 670 km. On the other hand, the effects would be greater if Sr extraction rates from the mantle were higher in the past (for example, because of higher continent growth rates), or if Sr from low-temperature sea-water alteration of basalt or from carbonates is recycled to the mantle. For example, Staudigel and Hart³⁸ estimated that $\sim 5 \times 10^{10}$ mol yr^{-1} of sea-water Sr are added to the oceanic crust by low-temperature alteration, which by itself equals half of the estimated continen-

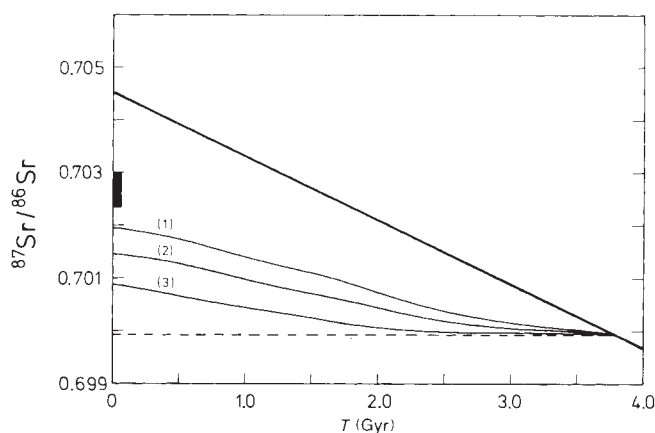


Fig. 5 Hydrothermal Sr exchange and $^{87}\text{Sr}/^{86}\text{Sr}$ in the upper mantle. We plot $^{87}\text{Sr}/^{86}\text{Sr}$ against time in the upper mantle from hydrothermal seawater Sr exchange with basalt, showing that this could significantly effect the $^{87}\text{Sr}/^{86}\text{Sr}$ of the upper mantle. The curves show only the effect of hydrothermal exchange (that is, as if there were no other source of ^{87}Sr). Calculations assumed that the upper mantle is 28% of the whole mantle³⁹, that there exists a uniform rate of depletion of Sr with time, and equimolar exchange between basalt and sea water at the rate of 10^{10} moles of Sr per 17 km^3 of oceanic crust per year (that is, present-day rates^{4,33}). Seawater $^{87}\text{Sr}/^{86}\text{Sr}$ are from refs 18 and 40. Curve 1 assumes that spreading rates are proportional to the square of the Earth's heat production⁴¹, and are therefore ~ 10 times the present-day values at 3.5 Gyr. Exchange at ridges would account for 3/4 of the increase in $^{87}\text{Sr}/^{86}\text{Sr}$ in the MORB source since 3.8 Gyr. For curve 2, which assumes spreading rates half that of curve 1 through time, exchange accounts for $\sim 1/2$ of the increase of $^{87}\text{Sr}/^{86}\text{Sr}$. Curve 3, assuming constant spreading rates at today's values, accounts for $\sim 1/3$ of the increase. The bold line indicates the bulk Earth evolution and the dashed line indicates the bulk Earth at 3.8 Gyr. The solid block indicates values for MORBs.

tal ^{87}Sr production by Rb decay. If this is recycled to the mantle, the estimate of recycled ^{87}Sr based on the hydrothermal flux is low by a factor of ~ 3.5 .

Preferential recycling of continental Sr to the mantle has implications for Nd and Sr isotopes in oceanic basalts. Estimates of the bulk Earth $^{87}\text{Sr}/^{86}\text{Sr}$ ratio based on Nd-Sr covariations in oceanic basalts have ranged from 0.7045 to 0.7052 (refs 42–44), and would be valid if ocean-island basalts with $\epsilon_{\text{Nd}} \approx 0$ (such as Tristan da Cunha, Gough and Koolau)^{42,45} have primitive sources. It has been shown that Nb/U and Ce/Pb in these lavas are very different from the values in primitive mantle but are like those in MORB^{46–48}, and therefore the sources are not primitive. Preferential recycling of continental Sr to the mantle shifts $^{87}\text{Sr}/^{86}\text{Sr}$ to higher values without affecting $^{143}\text{Nd}/^{144}\text{Nd}$. If oceanic island sources have been affected by addition of ^{87}Sr , then use of Nd-Sr isotope variations to infer Rb-Sr systematics of the bulk Earth is not justified. Independent estimates of bulk Earth Rb/Sr range from ~ 0.024 to 0.037, averaging ~ 0.03 , and imply that $^{87}\text{Sr}/^{86}\text{Sr} \approx 0.7036$ –0.7061 (refs 49–51).

Differentiation age of the mantle

If the Rb-Sr age of the whole continental crust is too young because it has lost ^{87}Sr , then this has been recycled to the convecting mantle or to the continental lithosphere. If the ^{87}Sr is in the convecting mantle, then its Rb-Sr differentiation age is too young. Estimates of the mean age of the continental crust are mostly > 2 Gyr (refs 6, 52–54), and some are > 2.5 Gyr (refs 4, 11). The Rb-Sr differentiation age of the MORB source (T_{MORB}) is younger, ~ 1.85 Gyr if $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{MORB}} \approx 0.7025$, $(\text{Rb}/\text{Sr})_{\text{nMORB}} \approx 0.1$, $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{BE}} = 0.7045$ and $(\text{Rb}/\text{Sr})_{\text{BE}} \approx 0.03$, where BE refers to the bulk Earth and n means normalized to the bulk Earth. T_{MORB} would be older if $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{BE}}$ or

$(\text{Rb}/\text{Sr})_{\text{MORB}}$ are higher, but younger if $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{MORB}}$ is higher⁵⁵.

If the MORB source represents the complement of the continental crust, then T_{MORB} could be the average age of continent-mantle differentiation. Whether inclusion of oceanic island sources would raise or lower the age is not clear; the higher $^{87}\text{Sr}/^{86}\text{Sr}$ by itself would lower it, but the higher Rb/Sr would raise it. Neither is it clear what would be a good estimate for $(\text{Rb}/\text{Sr})_{\text{OIB}}$, where OIB refers to the average value of the sources of oceanic island basalts, but this value is certainly lower than the Rb/Sr in the basalts. Primitive tholeiites have Rb/Sr most like their sources because they reflect a high degree of melting and minimal magmatic differentiation. Average $(\text{Rb}/\text{Sr})_{\text{n}}$ of 45 Icelandic tholeiites^{50,56} is 1.0, although the average of 7 primitive tholeiites is 0.26. For Hawaii, the average of 53 tholeiites and alkalic basalts^{42,57,58} is 0.81. Therefore, an upper limit is $(\text{Rb}/\text{Sr})_{\text{nOIB}} \approx 1$, although more reasonable estimates may be < 0.5 .

According to the crustal-residence-age equation, $T_a \leq T_b$ if $(1 + (I_a - I_s)/(P_a - P_s)) \leq [1 + (I_b - I_s)/(P_b - P_s)]$, where a and b are samples and s is the system. If $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{DM}}$ is known (where DM refers to the depleted portion of the convecting mantle, so that $\text{DM} \equiv \text{OIB}_s + \text{MORB}_s$), and the DM region adequately characterizes the mantle complement to the continental crust, then one can determine the $(\text{Rb}/\text{Sr})_{\text{OIB}}$ for which $T_{\text{DM}} = T_{\text{MORB}}$. It can be appreciated intuitively that if $T_{\text{DM}} = T_{\text{MORB}}$, a higher $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{DM}}$ requires higher $(\text{Rb}/\text{Sr})_{\text{OIB}}$. Also, a smaller OIB Sr reservoir requires higher $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{OIB}}$, and $(\text{Rb}/\text{Sr})_{\text{OIB}}$.

It can be shown that $T_{\text{DM}} \leq T_{\text{MORB}}$ if $P_{\text{OIB}_s} \leq (P_{\text{DM}} - P_{\text{MORB}_s})/f_{\text{OIB}_s}$, where P refers to the Rb/Sr ratio and f to the fraction of depleted-mantle Sr in the MORB and OIB sources. Therefore, if $(^{87}\text{Sr}/^{86}\text{Sr})_{\text{DM}} = 0.7030$ and $(\text{Rb}/\text{Sr})_{\text{nOIB}_s} \leq 1.0$ or 0.5, then $f_{\text{OIB}_s}/f_{\text{MORB}_s} \leq 0.32$ or 0.82, respectively. In either case a substantial portion of the Sr in the depleted mantle must be in the OIB reservoir before T_{DM} becomes older than T_{MORB} . If $(\text{Rb}/\text{Sr})_{\text{n}} \leq 0.25$, as suggested by the primitive Icelandic tholeiites, the size of the OIB reservoir would have to be impossibly large ($f_{\text{OIB}_s}/f_{\text{MORB}_s} > 100$) before $T_{\text{DM}} > T_{\text{MORB}}$. Therefore, it is likely that the Rb-Sr differentiation age of the whole depleted mantle is younger than the MORB differentiation age, and like the Rb-Sr age of the continental crust, it is too young.

Geological evidence

The realization that the Rb-Sr age of the continents is too young explains some enigmatic aspects of Sr isotopes in the continents. For example, the estimate of the growth rate of the continental crust by Hurley and Rand⁵⁹ is based primarily on Rb-Sr isotopes. They conclude that the rate of continental growth increased between 2.5 and 1.0 Gyr BP, with the maximum rate between 1.5 and 0.5 Gyr. Their estimate of the mean age is younger than 1.5 Gyr, which is much younger than estimates from later studies. They surveyed global Rb-Sr ages, compared initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios to those in the juvenile crust, and estimated a time interval for the radiogenic growth. Because initial $^{87}\text{Sr}/^{86}\text{Sr}$ values in continental rocks are too low, the method of Hurley and Rand inevitably results in too young an estimate. In general, because continental Sr is recycled preferentially to the mantle, precursor ages derived from initial $^{87}\text{Sr}/^{86}\text{Sr}$ in granite massifs to estimate maximum crustal ages (see, for example, ref. 60) are likely to be too young. If the Hurley and Rand⁵⁹ estimate accurately reflects the Rb-Sr age of the continents, then this is nearly 40% younger than the true age.

Conclusions

There is a contradiction between the high Rb/Sr in crustal rocks and an absence of a corresponding increase in initial $^{87}\text{Sr}/^{86}\text{Sr}$ in clastic sediments and other crustal rocks through time. Because $^{87}\text{Sr}/^{86}\text{Sr}$ values are too low, the Rb-Sr age of the

continental crust is too young, and younger than the Sm-Nd age. Sm and Nd remain in silicate rocks and are not significantly fractionated by common crustal processes, whereas Sr is leached during weathering. Through hydrothermal exchange of seawater Sr with basalt at spreading ridges, and probably through low-temperature seawater alteration of ocean-floor basalt and subduction of ocean-floor carbonates, continental Sr is lost

to the mantle preferentially to Nd. This has contributed significantly to the increase of $^{87}\text{Sr}/^{86}\text{Sr}$ in the upper mantle. As a result, the Rb-Sr differentiation age of the mantle is also too young.

This paper has benefitted from reviews and suggestions by F. Albarede, C. Chauvel, A. Hofmann, W. McDonough, H. Staudigel, P. Taylor and A. Zindler.

Received 30 August; accepted 7 November 1988.

- Garrels, R. M. & MacKenzie, F. T. *Evolution of the Sedimentary Rocks* 394 (Norton & Co., New York, 1971).
- Veizer, J. & Jansen, S. L. *J. Geol.* **87**, 341-370 (1979).
- Veizer, J. & Jansen, S. L. *J. Geol.* **93**, 625-643 (1985).
- Reymer, A. & Schubert, G. *Tectonics* **3**, 63-67 (1984).
- Taylor, S. R. *et al. Geochim. cosmochim. Acta* **47**, 1897-1905 (1983).
- Goldstein, S. L., O'Nions, R. K. & Hamilton, P. J. *Earth planet. Sci. Lett.* **70**, 221 (1984).
- O'Nions, R. K. *et al. Earth planet. Sci. Lett.* **63**, 229-240 (1983).
- Dasch, E. J. *Geochim. cosmochim. Acta* **33**, 1521-1556 (1969).
- Biscaye, P. & Dasch, E. J. *J. geophys. Res.* **76**, 5087-5096 (1971).
- Davies, G., Gledhill, A. & Hawkesworth, C. *Earth planet. Sci. Lett.* **75**, 1-12 (1985).
- Taylor, S. R. & McLennan, S. L. *The Continental Crust: Its Composition and Evolution* 312 (Blackwell Scientific Publications, London, 1985).
- Goldstein, S. J. & Jacobsen, S. B. *Earth planet. Sci. Lett.* **87**, 249-265 (1988).
- Fairbairn, H. W. *et al. Can. J. Earth planet. Sci. Lett.* **6**, 489-497 (1969).
- Obradovich, J. D. & Peterman, Z. E. *Can. J. Earth planet. Sci. Lett.* **5**, 737-747 (1968).
- Clauer, N. *Geochim. cosmochim. Acta* **37**, 2243-2255 (1973).
- Holland, H. D. *The Chemistry of the Atmosphere and Oceans* 351 (Wiley, New York, 1978).
- BVSP Basaltic Volcanism on the Terrestrial Planets (Pergamon, New York, 1981).
- Veizer, J. & Compston, W. *Geochim. cosmochim. Acta* **40**, 905-914 (1976).
- Gast, P. W. *J. geophys. Res.* **65**, 1287-1297 (1960).
- Peterman, Z. E. *et al. Earth planet. Sci. Lett.* **21**, 433-439 (1967).
- Allegre, C. J. & Ben Othman, D. *Nature* **286**, 335-342 (1980).
- McCulloch, M. T. & Chappell, B. W. *Earth planet. Sci. Lett.* **58**, 51-64 (1982).
- Ben Othman, D., Polvé, M. & Allège, C. J. *Nature* **307**, 510-515 (1984).
- Downes, H. & Leyreloup, A. *Geol. Soc. spec. Publ.* **24**, 319-330 (1986).
- Galer, S. J. G., Goldstein, S. L. & O'Nions, R. K. *Chem. Geol.* (in the press).
- Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7411-7427 (1979).
- DePaolo, D. J. *Geochim. cosmochim. Acta* **44**, 1185-1196 (1980).
- Edmond, J. M. *et al. Earth planet. Sci. Lett.* **46**, 1-18 (1979).
- Hart, S. R. & Staudigel, H. *Earth planet. Sci. Lett.* **58**, 202-212 (1982).
- Albarede, F. & Michard, A. *Chem. Geol.* **57**, 1-15 (1986).
- Spooner, E. T. C. *Earth planet. Sci. Lett.* **31**, 167-174 (1976).
- Brass, G. W. *Geochim. cosmochim. Acta* **40**, 721-730 (1976).
- Albarede, F. *et al. Earth planet. Sci. Lett.* **55**, 229-236 (1981).
- Cogley, J. G. *Rev. Geophys. Space Phys.* **22**, 101-112 (1984).
- O'Nions, R. K., Carter, S. R., Cohen, R. S., Evensen, N. M. & Hamilton, P. J. *Nature* **273**, 435-438 (1978).
- Javoy, M., Pineau, F. & Allège, C. J. *Nature* **300**, 171-173 (1982).
- Mattey, D. P. *Terra Cognita* **7**, 31-37 (1987).
- Staudigel, H. & Hart, S. R. *Geochim. cosmochim. Acta* **47**, 337-350 (1983).
- Dziewonski, A. M. & Anderson, D. L. (1981) *Phys. Earth planet. Inter.* **26**, 297-356 (1981).
- Burke, W. H. *et al. Geology* **10**, 516-519 (1982).
- Christensen, U. R. *J. geophys. Res.* **90**, 2995-3007 (1985).
- O'Nions, R. K. *et al. Earth planet. Sci. Lett.* **34**, 13-22 (1977).
- DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **31**, 743-746 (1976).
- Zindler, A., Jagoutz, E. & Goldstein, S. *Nature* **298**, 519-523 (1982).
- Stille, P., Unruh, D. M. & Tatsumoto, M. *Nature* **304**, 25-29 (1983).
- Newsom, H. E., White, W. M., Jochum, K. P. & Hofmann, A. W. *Earth planet. Sci. Lett.* **80**, 299-313 (1986).
- Hofmann, A. W., Jochum, K. P., Seufert, M. & White, W. M. *Earth planet. Sci. Lett.* **79**, 33-45 (1986).
- Hofmann, A. W. *Chem. Geol.* **57**, 17-30 (1986).
- Anderson, D. L. *J. geophys. Res.* (Proc. 14th Lunar planet. Sci. Conf.) **88**, B41-B52 (1983).
- Wood, D. A. *et al. Contr. Miner. Petrol.* **70**, 319-339 (1979).
- Sun, S. S. *Geochim. cosmochim. Acta* **46**, 179-192 (1982).
- Allège, C. J. & D. Rousseau, D. *Earth planet. Sci. Lett.* **67**, 19-34 (1984).
- Michard, A., Gurriet, P., Soudant, M. & Albarede, F. *Geochim. cosmochim. Acta* **49**, 601-610 (1985).
- Miller, R. G. *et al. Chem. Geol.* **57**, 87-99 (1986).
- White, W. M., Hofmann, A. W. & Puchelt, H. *J. geophys. Res.* **92**, 4881-4893 (1987).
- Oskarsson, N. *et al. J. Petrol.* **23**, 28-74 (1982).
- Chen, C. Y. Frey, F. A. *J. geophys. Res.* **90**, 8743-8768 (1985).
- Hofmann, A. W. *et al. Contr. Miner. Petrol.* **88**, 24-35 (1984).
- Hurley, P. M. & Rand, J. R. *Science* **164**, 1229-1242 (1969).
- Vidal, P. *Mém. Soc. géol. minér. Bretagne* **21**, 162 (1980).
- Allsopp, H. L. & Kolbe, P. *Geochim. cosmochim. Acta* **29**, 1115-1130 (1965).
- Andre, L., Deutsch, S., & Hertogen, J. *Chem. Geol.* **27**, 101-115 (1986).
- Stosch, H. G. & Lugmair, G. W. *Nature* **311**, 368-370 (1984).
- Whitney, P. R. & Hurley, P. M. *Geochim. cosmochim. Acta* **28**, 425-436 (1984).
- Gebauer, D. & Gruenenfelder, M. *Contr. Miner. Petrol.* **47**, 113-130 (1974).
- McCulloch, M. T. & Wasserburg, G. J. *Science* **200**, 1003-1011 (1978).
- Jacobsen, S. B. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **50**, 139-155 (1980).
- White, W. M. *Geology* **13**, 115-118 (1985).
- Weaver, B. & Tarney, J. *Nature* **310**, 575-578 (1984).
- Shaw, H. *et al. Geol. Soc. Spec. Publ.* **24**, 275-282 (1986).
- Wedepohl, K. H. *Fortschr. Miner.* **52**, 141-172 (1975).
- Martin, J. M. & Meybeck, M. *Mar. Chem.* **7**, 173-206 (1979).
- Wedepohl, K. H. *Fortschr. Miner.* **59**, 203-204 (1981).

Human cdc2 protein kinase is a major cell-cycle regulated tyrosine kinase substrate

Giulio Draetta*, Helen Piwnicka-Worms†, Deborah Morrison†, Brian Druker†, Thomas Roberts† & David Beach*

* Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, New York 11724, USA

† Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA

cdc2 is a catalytic subunit of a protein kinase complex, called the M-phase promoting factor, that induces entry into mitosis and is universal among eukaryotes. In HeLa cells, *cdc2* is shown to be the most abundant phosphotyrosine-containing protein and its phosphotyrosine content is subject to cell-cycle regulation. One site of *cdc2* tyrosine phosphorylation in vivo is selectively phosphorylated by pp60^{c-src} in vitro.

THE pathways that regulate eukaryotic cell proliferation have been investigated from two different perspectives. One approach, usually known as the study of 'growth control', has been to characterize the cellular and biochemical events that are associated either with mitogenic activation of non-dividing cells by growth factors or with transformation by oncogenic viruses. The other route has been to identify regulators of the division cycle in cells that are synchronized from actively growing cultures, or that divide with natural synchrony. A key criterion in assessing the role of an element that might be involved in 'cell-cycle control' is whether it behaves as a cell-cycle oscillator: whether its biochemical activity or some associated property fluctuates during the division cycle.

Among the numerous genes that are transcriptionally activated during mitogenesis, some such as *fos* are transcribed only transiently and others such as *myc* are subsequently expressed at constant levels in continuous culture^{1,2}. It has therefore been argued that the events of mitogenic activation may have little relevance in continuously dividing cells. The activators of cell growth however, are expected to influence the critical oscillators that regulate passage through phases of the division cycle. In this study, we identify a biochemical interconnection between growth control and cell-cycle control. We show that the human *cdc2* protein kinase, a component of a regulatory cell-cycle oscillator, is the major phosphotyrosine-containing polypeptide in proliferating HeLa cells, and that its phosphotyrosine content

increases during the progression from G1 to mitosis. Furthermore, one site of cdc2 tyrosine phosphorylation *in vivo* is phosphorylated by pp60^{c-src} *in vitro*, suggesting that this tyrosine kinase, or one with similar specificity, contributes to regulation of the cell cycle.

Tyrosine kinases

Many viral oncogenes and growth-factor receptors encode tyrosine kinases, and these fall broadly into two classes. One is the cell-surface receptors of growth factors such as insulin, platelet-derived growth factor, epidermal growth factor and, colony-stimulating factor. Each receptor is a membrane-spanning protein that has an extracellular ligand-binding domain. Binding of the growth factor to the respective extracellular domain activates the tyrosine kinase activity of the receptor and the ensuing phosphorylation of critical substrates is presumed to trigger the mitogenic response³. The second group, typified by pp60^{c-src}, lacks an extracellular domain and instead resides on the inner surface of the plasma membrane. Several members of both classes of tyrosine kinases have been transduced into oncogenic viruses³.

The cdc2 mitotic oscillator

The best characterized cell-cycle regulator in eukaryotic cells is a protein kinase of relative molecular mass (M_r) 34,000 (34 K) encoded by the *cdc2/CDC28* gene. This protein kinase was identified as the product of the *CDC28* cell-cycle 'start' gene of the budding yeast *Saccharomyces cerevisiae*⁴⁻⁶. Subsequently, *CDC28* was shown to be the homologue of the fission yeast *Schizosaccharomyces pombe* *cdc2* gene⁷, and a homologous protein kinase has now been found in every eukaryotic species that has been investigated, including man⁸⁻¹¹. In human HeLa cells, the cdc2 protein kinase exists in a complex with two other polypeptides. One, (13K) is the homologue of the *Schiz. pombe* *suc1* gene^{9,12}, and the other is a 62K species that has yet to be identified genetically¹³. The abundance of cdc2 remains constant throughout the cell cycle in human cells but its activity oscillates dramatically¹³. Kinase activity is barely detectable early in the G1 phase and is maximal during mitotic metaphase. Activation is associated with both cell-cycle dependent phosphorylation and also rearrangement of the subunit composition of the protein kinase complex.

In *Sacc. cerevisiae*, CDC28 contributes to regulation of the division cycle at the G1 'start' point^{4,5}, but in fission yeast *cdc2* is required not only for the initiation of nuclear DNA synthesis but also for entry into mitosis¹⁴. In HeLa cells, the pattern of activity of the kinase during the cell cycle indicates a possible involvement in the regulation of entry into mitosis¹³. This suggestion is strongly supported by the demonstration that the *Xenopus laevis* homologue of the cdc2 protein is a component of the M-phase promoting factor (MPF)^{15,16}, an activity that is assayed by its ability to drive interphase nuclei into mitosis^{17,18}. *cdc2* is also a component of the M-phase specific histone-H1 kinase, which is probably the same entity as MPF^{19,20}. Two further cdc2-associated proteins, at least in clams, are cyclin A and B²¹. These proteins are specifically proteolysed at the metaphase/anaphase transition of mitosis²², a process that seems to be responsible for the inactivation of MPF²¹. It is not known at present whether human p62 is a cyclin but the known properties of the protein are compatible with this possibility¹³. Taken together, these observations indicate that in multicellular organisms cdc2 acts at a critical step in the initiation of mitosis. The appealing possibility that this protein kinase activity might also be involved in the initiation of DNA synthesis in vertebrate cells has yet to be investigated.

cdc2 phosphotyrosine

Immunoprecipitation of cdc2 from a lysate of ³⁵S labelled asynchronous HeLa cells reveals a 34K protein (p34) and two associated polypeptides, p13 and p62 (ref. 13, Fig. 1a). p34 resolves

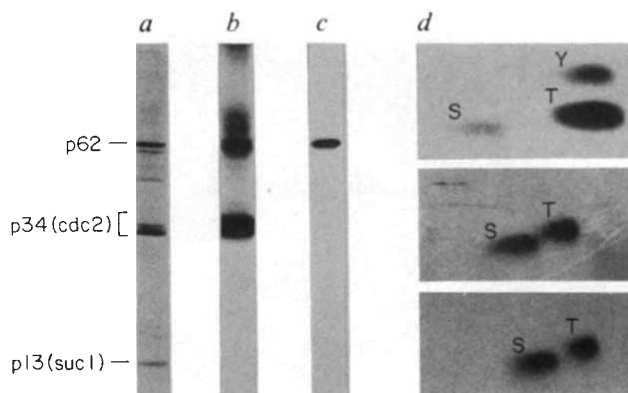


Fig. 1 Phosphoamino-acid analysis of components of the cdc2 complex. **a**, cdc2 complex immunoprecipitated from ³⁵S-labelled HeLa cell lysate. **b**, *in vivo* ³²P-labelled cdc2 complex. **c**, *In vitro* phosphorylation of p62 in the immuno-complex. **d**, 2-D-phosphoamino-acid analysis: from top to bottom, p34 (*in vivo*), p62 (*in vivo*) and p62 (*in vitro*). S, serine; T, threonine; Y, tyrosine.

Methods. Cell culture, labelling, immunoprecipitation with anti-cdc2 serum and kinase assays were as previously described¹⁰. For the phosphoamino-acid analysis ³²P-labelled protein bands were excised from dried polyacrylamide gels and proteins were extracted as described³⁹. The protein pellet was digested for 1.5 h at 110 °C in constant boiling HCl (Pierce). After lyophilization, the hydrolysate was resuspended in pH 1.9 electrophoresis buffer [88% formic acid/glacial acetic acid/water, 50:156:1794 (vol/vol)] and applied to a 0.1 mm thin-layer cellulose plate (Eastman Kodak). The labelled samples and a mixture of unlabelled phosphoamino-acid standards were separated by electrophoresis at 800 V for 30 min. After air-drying the plates, samples were analysed by ascending chromatography in a solution composed of isopropanol/hydrochloric acid/water (7:1.5:1.5). The position of unlabelled standards was determined by ninhydrin staining; labelled amino acids were detected by autoradiography.

into three bands of slightly different electrophoretic mobility, which have previously been shown to represent different phosphorylation states of the protein¹³. After incubation of HeLa cells with inorganic ³²P, both p34 and p62 but not p13 become labelled (Fig. 1b). If instead, the cdc2 complex is immunoprecipitated from unlabelled cell lysates and incubated with γ -[³²P]ATP *in vitro*, only p62 becomes phosphorylated¹³ (Fig. 1c). At present, we assume that p62 is a substrate of p34, but the possibility that p62 is also a protein kinase has not been rigorously excluded.

Phosphoamino-acid analysis of p34 and p62 labelled *in vivo* and p62 labelled *in vitro* was undertaken. p34 contained predominantly phosphothreonine, but also a significant fraction of phosphotyrosine and lower levels of phosphoserine (Fig. 1d, upper panel). In contrast, p62, labelled either *in vivo* or *in vitro*, contained roughly equal amounts of phosphoserine and phosphothreonine (Fig. 1d, middle and lower panels respectively) but no phosphotyrosine was detected. These observations indicate that HeLa cells contain a tyrosine kinase of which the cdc2 protein is a substrate.

Cell-cycle regulation

The overall level of phosphorylation of cdc2 in HeLa cells is subject to cell-cycle regulation¹³. We have investigated whether the tyrosine component of this phosphorylation is regulated or constant during the division cycle. HeLa cells were synchronized by centrifugal elutriation, which separates cells according to sedimentation velocity and thus according to size. Cell size broadly reflects cell-cycle position (Fig. 2a). The size-fractionated cells were lysed, part of the lysate was used for anti-p34 immunoblotting (Fig. 2c) and the remainder was immunoprecipitated with anti-p34 serum. A fraction of each immunoprecipitate was incubated with histone H1 as an

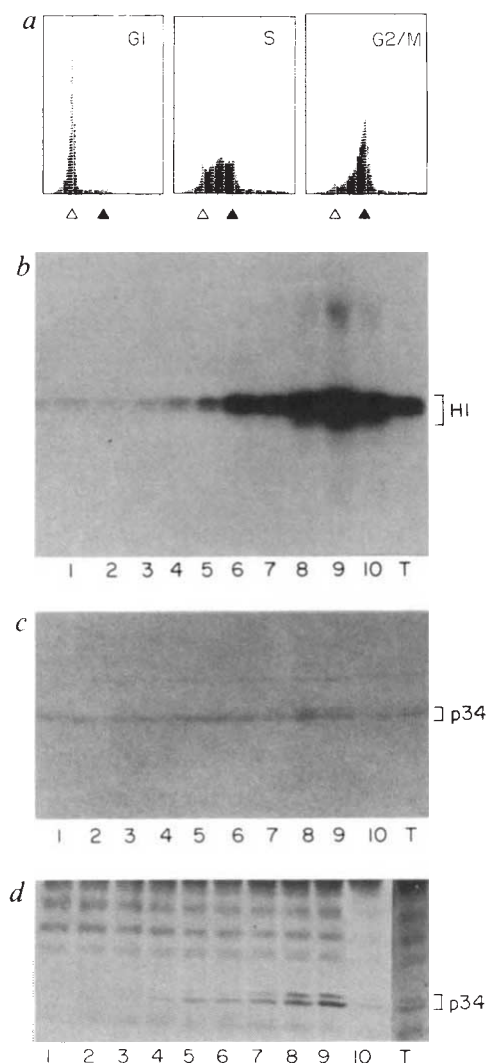


Fig. 2 Cell-cycle regulation of *cdc2* activity and tyrosine phosphorylation. *a*, Flow-cytometric profiles of HeLa cells separated by centrifugal elutriation. The G1, S and G2/M enriched fractions are numbers 1, 6 and 9 respectively. Open and closed triangles mark 2N and 4N DNA content, respectively. *b*, Protein kinase assay of the *cdc2* complex prepared from lysates of elutriated cell fractions. Samples were normalized for protein content before anti-*cdc2* peptide immunoprecipitation. Histone-H1 was the substrate. *c*, Immunoblot of 100 μ g of lysate from each fraction, probed with anti-*cdc2* serum. *d*, Anti-phosphotyrosine immunoblot of material precipitated from elutriated cell lysates with anti-*cdc2* peptide serum. Numbers below each fraction indicate the elutriated fraction number. (T) is the sample taken before elutriation.

Methods. Centrifugal elutriation and flow cytometric analysis of HeLa cells were performed as described¹³. For protein kinase assays¹³, 0.05 mg ml⁻¹ histone H1 (Boehringer) was used as substrate. Immunoblotting was done using anti p34^{cdc2} polyclonal serum¹⁰. Immunoblotting with anti-phosphotyrosine antibodies was as described²³, using a serum prepared against phosphotyrosine coupled to keyhole limpet haemocyanin and affinity-purified on a phosphotyramine-Sepharose column. Immunological reactions were visualized with an alkaline phosphatase detection system.

The anti-phosphotyrosine immunoblot of the anti-p34 immunoprecipitation revealed two specific bands (Fig. 2*d*). These bands displayed a slightly retarded electrophoretic mobility with respect to unphosphorylated fission yeast p34 purified from *Escherichia coli*¹⁰ (data not shown), and hence represent the two phosphorylated forms of the three forms of p34¹³. The intensely labelled bands migrating with lesser electrophoretic mobility than p34 are due to cross-reaction between the heavy chain of the antibody used to immunoprecipitate the *cdc2* complex and the secondary antibody used in immunoblotting. The level of tyrosine phosphorylation of *cdc2* showed a clear cell-cycle fluctuation that paralleled the biochemical activity of the protein kinase. We do not, however, presume from this data that p34 is most heavily tyrosine-phosphorylated at the time of maximal activity during mitosis, as this relatively brief cell-cycle phase is not clearly resolved by elutriation. These observations do suggest either that the tyrosine kinase that phosphorylates p34 is regulated during the cell cycle or that there exists a cell-cycle dependent phosphotyrosine phosphatase.

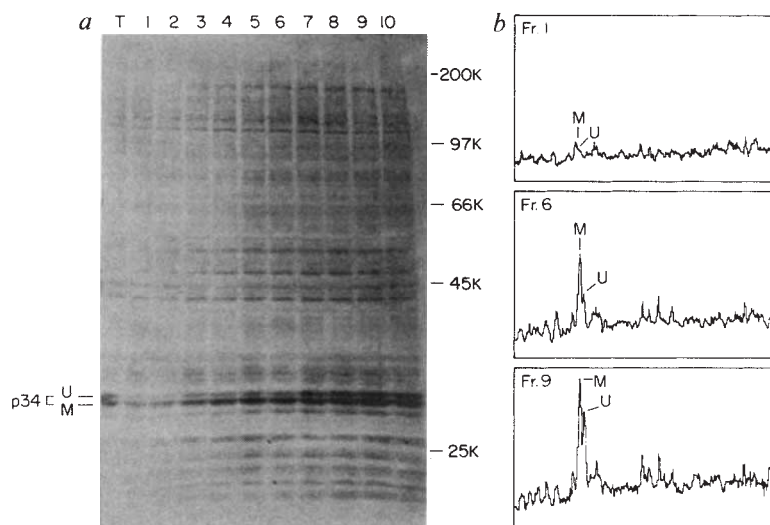
To test whether *cdc2* is an abundant or minor phosphotyrosine containing protein in HeLa cells, the elutriation experiment was repeated but instead of immunoprecipitating p34 before blotting, total cell protein was resolved by SDS-PAGE and immunoblotted with anti-phosphotyrosine antibodies (Fig. 3). Remarkably, among the most intensely staining bands in unsynchronized cells were a doublet of the same mobility as the two phosphotyrosine-containing forms of p34 (Fig. 3*a*, lane T). These two bands also displayed the same pattern of cell-cycle fluctuation as *bona fide* p34 (Figs 2*d* and 3*a*). Thus, during progression from G1 towards mitosis, the overall level of phosphotyrosine increased in this polypeptide, as did the ratio of the most phosphorylated (upper band, designated U) to the intermediate form of *cdc2* (lower band, designated M, Fig. 3*b*). It should be noted that the slower migration of these bands is not necessarily due to phosphotyrosine, and there is no way of knowing whether p34 is phosphorylated at more than one tyrosine residue from these data.

Two experiments were undertaken to test formally whether this abundant tyrosine-phosphorylated protein is indeed *cdc2*. First, a total HeLa cell lysate was precleared with anti-p34 antiserum that was covalently coupled to Sepharose (see Fig. 4 legend). After electrophoresis of the lysate, the immunoblot was probed with anti-phosphotyrosine serum. Compared with an untreated cell extract, the precleared preparation displayed selective loss of the major 34K phosphotyrosine-containing species (Fig. 4*a*). In a further test, we used a quite different anti-phosphotyrosine serum. That used the previous experiments was prepared by immunizing rabbits with phosphotyramine coupled to keyhole limpet haemocyanin²³. The second antibody was obtained from serum of rabbits immunized with autophosphorylated v-abl that was subsequently affinity-purified on a phosphotyramine column. The previous experiment was repeated with this serum. We confirmed that the preclearing procedure totally depleted *cdc2* (Fig. 4*b*, left panel) and found that the 34K species, which was by far the most reactive of all cellular proteins with the second anti-phosphotyrosine serum, was likewise selectively depleted (Fig. 4*b*, right panel). We conclude from these data that p34 is one of the most abundant phosphotyrosine-containing proteins in HeLa cells and that its level of tyrosine phosphorylation is subject to cell-cycle regulation.

HeLa cells contain somewhat higher levels of p34 than some other types. For example, we have estimated that p34 represents 0.05% of total HeLa cell protein, a fraction very similar to that in *Xenopus* eggs¹⁵. Actively dividing mouse NIH3T3 cells have approximately half this level and in other cell types the level of p34 is highly variable (G.D. and D.B., unpublished results). In the light of the relatively high level of phosphotyrosine in p34, it is important to bear in mind that although HeLa cells are highly transformed, they do not display the levels of total cell

exogenous substrate in the presence of [γ -³²P]ATP (Fig. 2*b*) and the remainder was analysed by SDS-PAGE followed by immunoblotting with a polyclonal anti-phosphotyrosine serum²³ (Fig. 2*d*). In this experiment, the level of p34 remained approximately constant (Fig. 2*c*), but the kinase activity of the p34 complex varied 70-fold across the elutriation gradient (Fig. 2*b*).

Fig. 3 Anti-phosphotyrosine immunoblot of whole cell lysates of elutriated cells. *a*, 100 μ g of HeLa cell lysate from each elutriated fraction were resolved by SDS-PAGE and transferred to nitrocellulose filter. Protein blots were probed with anti-phosphotyrosine antibodies²³. Relative molecular mass markers are shown at the side. *b*, Laser-densitometer scanning of lane 1, 6 and 9 from the immunoblot in *a*. U and M indicate the upper and middle bands of the p34 triplet.



phosphotyrosine that are found in cells transformed by oncogenes encoding tyrosine kinases²³.

In vitro substrate of pp60^{c-src}

Among the known tyrosine kinases, only one has been shown to be subject to cell-cycle regulation. In a mouse fibroblast NIH3T3 cell line that overexpresses pp60^{c-src} by approximately 20-fold, pp60^{c-src} is several times more active in mitosis than during interphase²⁴. We therefore investigated at the ability of pp60^{c-src} to phosphorylate purified p34 *in vitro*. Using standard biochemical techniques p34 was purified from HeLa cells (Fig. 5a). Western blotting was used to track the p34 in the early

stages of column chromatography and either Coomassie blue or silver staining at later steps. The material obtained in this manner was not associated with p62, retained the p13 subunit, but lacked detectable protein kinase activity. Purified p34 was incubated with pp60^{c-src} that had been immunoprecipitated from insect cells infected with a recombinant baculovirus encoding avian c-src (see Fig. 5 legend).

In the absence of added p34, only pp60^{c-src} autophosphorylation was observed (Fig. 5b, lane 1). However, in the presence of p34 at 10⁻⁸ M, phosphorylation of a polypeptide of 34K also occurred (lane 2). In immunoprecipitations of insect cells infected with a virus encoding pp60^{c-src} Met295, a mutant that lacks protein kinase activity²⁵, no phosphorylation of either pp60 or the 34K species could be detected (Fig. 5b, lanes 3,4). Thus, phosphorylation of the 34K protein is dependent on the kinase activity of pp60^{c-src}. To establish formally that the 34K c-src substrate is the cdc2 protein, rather than a minor contaminant in the protein preparation, *in vitro*-labelled 34K material was co-electrophoresed with an excess of recombinant fission yeast p34 that had been expressed and purified from *E. coli*¹⁰. After Coomassie blue staining, p34 was excised from the gel and exposed to *N*-chlorosuccinimide, a reagent that specifically cleaves polypeptides at tryptophan residues under appropriate conditions²⁶. The partially cleaved material in the gel slice was subjected to further gel electrophoresis and the bacterial p34 was visualized by silver staining whereas the 34K protein labelled *in vitro* was detected by autoradiography. The pattern of ³²P and silver-staining bands was virtually identical (Fig. 5c). As there is complete conservation of the number and distribution of tryptophan residues in human and fission yeast p34^{10,11,27} this result confirms that pp60^{c-src} phosphorylated authentic HeLa p34 *in vitro*.

pp60^{c-src} might phosphorylate proteins *in vitro* that are not necessarily substrates *in vivo*. If, however, pp60^{c-src} is the physiologically relevant tyrosine kinase in the cell, it is anticipated that the tyrosine residue(s) on which p34 is phosphorylated *in vivo* will be the same as that used by pp60^{c-src} *in vitro*. To this end, we performed comparative two-dimensional tryptic phosphopeptide mapping (Fig. 6). p34 was labelled *in vivo* with [³²P]-labelled inorganic phosphate and immunoprecipitated as described in Fig. 1b. Alternatively, purified p34 was labelled *in vitro* by incubation with pp60^{c-src}. Labelled p34 was digested with trypsin, and the resulting phosphopeptides were separated by electrophoresis, at pH 8.9 or 1.9, in the first dimension and chromatography in the second dimension. Trypsin digestion of p34 labelled *in vivo* generated three major phosphopeptides designated A, B and C (Fig. 6, *a*, *d*). p34 labelled *in vitro* gave rise to one major phosphopeptide designated 1 (Fig. 6, panels

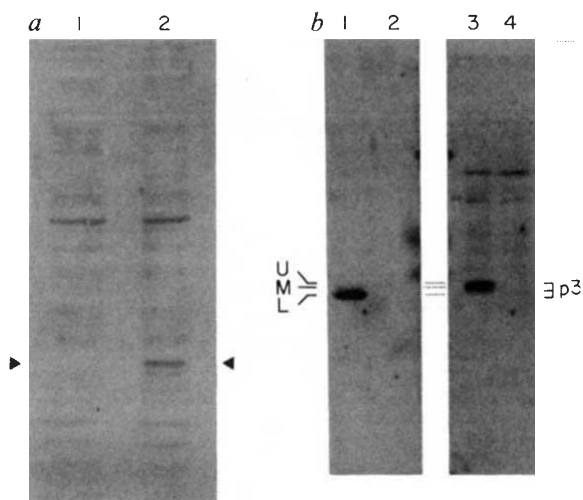
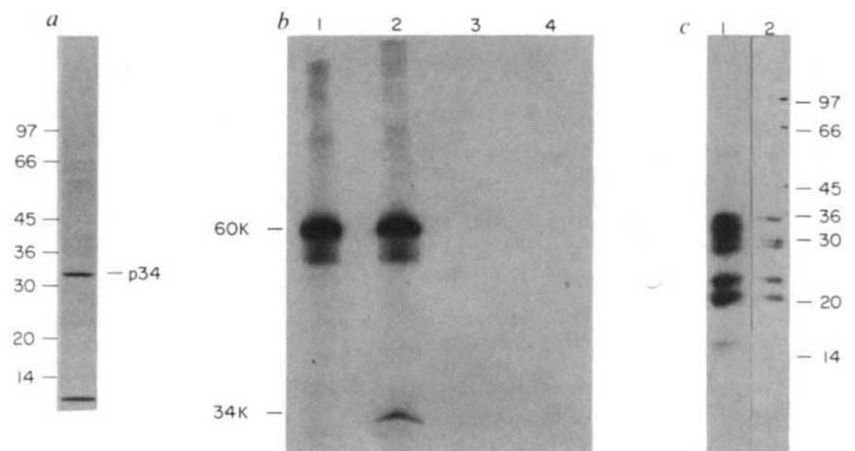


Fig. 4 Anti-phosphotyrosine immunoblots of cell lysates cleared of cdc2. *a*, 100 μ g of HeLa cell lysate was reacted for four hours with anti-cdc2 coupled to Sepharose and the unbound material was subjected to blotting with anti-phosphotyrosine serum (lane 1). Lane 2 shows a sample that was precleared with control Sepharose. The diamonds indicate the position of the 34K protein that is selectively depleted. *b*, Lane 1. Anti-cdc2 immunoblot of HeLa whole-cell lysate. Lane 2, anti-cdc2 immunoblot of lysate precleared with anti-cdc2 Sepharose. U, M and L indicate the three forms of p34. The lower, unphosphorylated band, is predominant in this blot. Lanes 3 and 4 represent an exact duplication of lanes 1 and 2, respectively, except that this half of the immunoblot was probed with anti-phosphotyrosine serum that had been prepared from rabbits immunized with autophosphorylated v-abl. The lines between the two blots indicate precise alignment of the two halves of a single gel.

Fig. 5 Phosphorylation of *cdc2* by $\text{pp60}^{\text{c-src}}$ *in vitro*. **a**, Silver staining SDS-PAGE of *cdc2* purified from HeLa cells. **b**, Phosphorylation of p34 by recombinant $\text{pp60}^{\text{c-src}}$: $\text{pp60}^{\text{c-src}}$ immunoprecipitation and kinase assay in the absence (lane 1) or in the presence (lane 2) of 10 nM p34; $\text{pp60}^{\text{c-src}}$ Met295 immunoprecipitation and kinase assay in the presence (lane 3) or the absence (lane 4) of p34. **c**, *N*-chlorosuccinimide digestion of mixture of unlabelled fission yeast p34 and human p34 phosphorylated by $\text{pp60}^{\text{c-src}}$: lane 1, autoradiogram; lane 2, silver staining of the digested material which was largely fission yeast p34.

Methods. Purification of p34 from HeLa cells. Immunoblotting was used to detect p34 during the early stages of column fractionation. Eight litres of HeLa cells at a density of 4×10^5 – $6 \times 10^5 \text{ ml}^{-1}$ were centrifuged at 1,500g for 10 min, washed in PBS and resuspended in 40 ml buffer 1 (10 mM Hepes, pH 7.4, 10 mM NaCl and 0.1 mM EDTA). Protease inhibitors were added as previously described¹⁰, and cell extract was prepared using a Dounce homogenizer (40 strokes). After centrifugation for 1 h at 100,000g, the supernatant was applied to a 40 ml DE52 (Whatman) column (2.5 × 8 cm) equilibrated in buffer 2 (50 mM HEPES-NaOH, pH 7.4, containing 5 mM MgCl_2 , 1 mM EDTA, 1 mM DTT). The flow-through fraction and two 40-ml buffer 2 washes were collected, pooled and applied to a 10 cm (1 cm × 12 cm) column of Sepharose S FF (Pharmacia) at a flow rate of 0.8 ml min^{-1} . After a 20-ml wash, a gradient of 65 ml of buffer 2 and 65 ml buffer 3 (buffer 2 containing 0.5 M NaCl) was applied and 2.5-ml fractions were collected, p34-containing fractions were pooled and loaded onto a 1-ml hydroxylapatite column equilibrated in buffer 4 (20 mM Tris-HCl, pH 7.4, 5 mM MgCl_2 , 0.1 M NaCl, 1 mM EDTA, 0.2 mM dithiothreitol, DTT, 10% glycerol). The flow rate was set at 4 ml h^{-1} . After a 5-ml wash, a gradient of 10 ml buffer 5 (0.01 M potassium phosphate, pH 7.5, 0.1 M NaCl, 10% glycerol, 5 mM MgCl_2 , 1 mM EDTA, 0.2 mM DTT) and 10 ml buffer 6 (buffer 5 containing 0.15 M potassium phosphate) was applied and 0.5 ml fractions were collected. Fractions containing p34 were pooled, dialysed against buffer 7 (10 mM HEPES-NaOH, pH 7.4, containing 5 mM MgCl_2 , 1 mM EDTA, 0.2 mM DTT, 10 mM NaCl, 10% glycerol) and applied to a 1-ml casein-Sepharose column. After a 5 ml wash, a gradient of 10 ml buffer 7 and 10 ml buffer 8 (buffer 7 containing 0.5 M NaCl) was applied and 0.5 ml fractions were collected at a flow rate of 2.5 ml h^{-1} . Fractions were then analysed by SDS-PAGE and silver staining. p34 containing fractions were concentrated against buffer 7 containing 50% glycerol and stored at -70°C . Approximately 10 μg of p34 monomer can be purified with an estimated recovery of approximately 10% from eight litres of HeLa cells.

cDNAs encoding avian $\text{pp60}^{\text{c-src}}$ ⁵⁰ and a *c-src* mutant allele encoding a methionine rather than a lysine residue at position 295 (ref. 35) were cloned into a baculovirus expression vector⁴¹. Recombinant viruses encoding $\text{pp60}^{\text{c-src}}$ and $\text{pp60}^{\text{c-src}}$ Met295 were plaque purified and used to infect *Spondoptera frugiperla* (SF9) cells. All procedures relating to viral propagation, isolation and infection performed essentially as described⁴¹. $\text{pp60}^{\text{c-src}}$ and $\text{pp60}^{\text{c-src}}$ Met295 comprised approximately 1% of total cell protein, each in a soluble form (Piwnicka-Worms *et al.*, unpublished data). Forty hours after infection of SF9 cells with recombinant virus encoding either $\text{pp60}^{\text{c-src}}$ or $\text{pp60}^{\text{c-src}}$ Met²⁹⁵, cell lysates were prepared and immunoprecipitated with EC10, a monoclonal antibody specific for $\text{pp60}^{\text{c-src}}$ ⁵². The immunoprecipitates were then incubated in either the presence or absence of p34 and kinase assays were performed *in vitro*. *N*-chlorosuccinimide cleavage of p34 was performed exactly as previously described¹⁰.



b, *e*) and at pH 1.9 a minor species designated 2 (Fig. 6, *e*).

To determine whether the site of p34 phosphorylation *in vitro* might correspond to one of the three major sites *in vivo*, we mixed tryptic digestion products from *in vivo* and *in vitro* labelling (Fig. 6, *c* and *f*). The *in vivo* labelled phosphopeptide designated C co-migrated with the major peptide phosphorylated by $\text{pp60}^{\text{c-src}}$ *in vitro*. The co-migration of these peptides, at both pH 8.9 and 1.9, provides strong evidence that species 1 and C are the same peptide.

We performed phosphoamino-acid analysis to test whether peptide C contained phosphotyrosine, as would be expected if it contains a site of *in vivo* $\text{pp60}^{\text{c-src}}$ phosphorylation. One-dimensional phosphoamino-acid analysis was undertaken on each phosphopeptide recovered by elution from the two-dimensional chromatographs. Undigested p34, labelled *in vivo*, contained phosphothreonine, phosphotyrosine and phosphoserine, whereas p34 labelled *in vitro* by $\text{pp60}^{\text{c-src}}$ contained only phosphotyrosine (Fig. 7). Tryptic peptide A contained phosphothreonine and some phosphotyrosine, whereas species B contained only phosphothreonine. Peptide C, the equivalent of peptide 1 labelled by $\text{pp60}^{\text{c-src}}$ *in vitro*, contained only phosphotyrosine. These results suggest that $\text{pp60}^{\text{c-src}}$ is capable of phosphorylating *cdc2* *in vivo* at a site used by a tyrosine kinase *in vivo*.

Discussion

We have shown that, in addition to phosphothreonine and phosphoserine, human p34 contains phosphotyrosine. Indeed, p34 is among the most abundant phosphotyrosine-containing polypeptides in HeLa cells and the extent of this post-transla-

tional tyrosine phosphorylation varies during the cell cycle. *In vitro*, $\text{pp60}^{\text{c-src}}$ is able to phosphorylate p34. A site within p34 that is recognized by $\text{pp60}^{\text{c-src}}$ *in vitro* is one of the two at which p34 is phosphorylated on tyrosine *in vivo*. This suggests either that p34 is a substrate of $\text{pp60}^{\text{c-src}}$ *in vivo*, or that the physiologically relevant tyrosine kinase has a similar ability to phosphorylate p34. The demonstration that two of the three tryptic phosphopeptides of p34 labelled *in vivo* contain phosphotyrosine, whereas $\text{pp60}^{\text{c-src}}$ phosphorylates only one predominant site *in vitro*, may suggest that p34 is a substrate of more than one tyrosine kinase *in vivo*.

These observations raise various questions about the role of tyrosine phosphorylation in the activation of the *cdc2* kinase complex and in cell-cycle control. It should be borne in mind that in *Xenopus* the *cdc2* kinase is component of MPF. MPF is a non-species-specific factor that acts at the G2/M boundary to initiate a cascade of events taking the interphase cell into mitosis^{17,18}. These include dissolution of the cytoplasmic network of microtubules, formation of the mitotic spindle, breakdown of the nuclear membrane and mitotic chromosome condensation. Most of these events have been reconstructed in cell-free lysates and can be directly triggered by addition of MPF²⁸. It has been known for many years that mitotic HeLa cells have an MPF activity that may be assayed by microinjection into *Xenopus* oocytes²⁹. The biochemical properties of the *cdc2* mitotic complex, particularly its activation during mitosis, are consistent with the possibility that it represents the human MPF¹³ but this has not been formally proved.

Human *cdc2* is activated in multiple steps during the transition from G1 to mitosis¹³. Not only is p34 progressively phosphory-

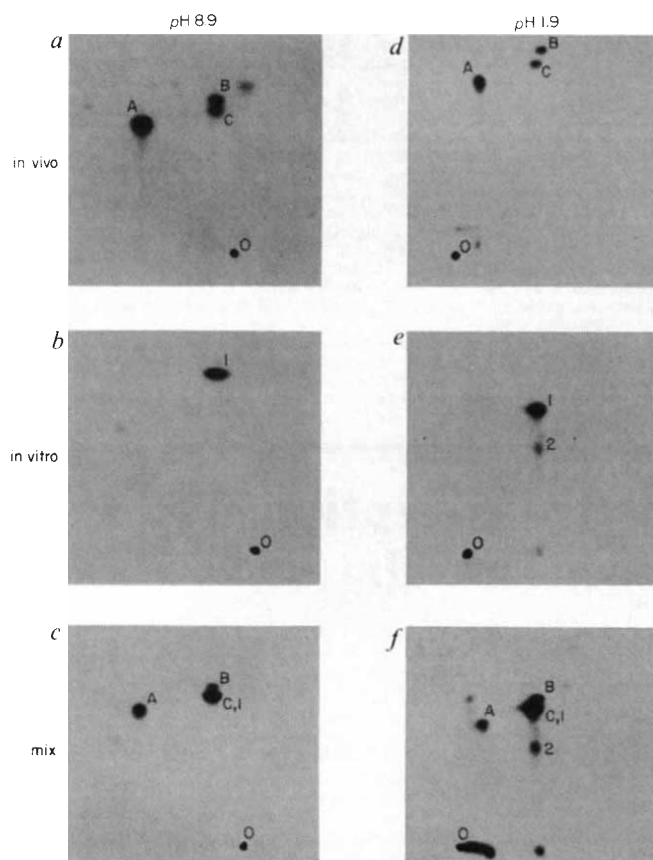


Fig. 6 Two-dimensional tryptic phosphopeptide analysis of *cdc2* labelled *in vivo* (a, d), or *in vitro* by pp60^{c-src} (b, e) or a mixture of both (c, f). The electrophoretic dimension was performed at either pH 8.9 or 1.9.

Methods. p34 immunoprecipitated from *in vivo* ³²P-labelled cells, or purified p34 phosphorylated *in vitro* by pp60^{c-src} were resolved by SDS-PAGE. Labelled p34 was eluted from the gel, digested with trypsin and the resulting phosphopeptides were separated by electrophoresis (at pH 1.9 or 8.9) in the first dimension (horizontal) and chromatography in the second dimension (vertical) as described⁴². The origin is marked with an O.

lated during the cell cycle but the most phosphorylated form also associates with the p62 subunit which is thought to have a regulatory role¹³. The level of tyrosine phosphorylation increases during the division cycle and probably contributes to the activating process. At present, however, it is impossible to dissociate the effect of p34 tyrosine phosphorylation from the other events such as threonine phosphorylation of the protein and complex formation with p62. It might be relevant that pp60^{v-src}, which is more active than pp60^{c-src} but has a similar substrate specificity³⁰, can accelerate the process of progesterone-induced meiotic maturation upon microinjection into *Xenopus* oocytes³¹. Of course, a key step in *Xenopus* oocyte maturation is the elaboration of active MPF, of which *cdc2* is one component.

More generally, the phosphorylation of *cdc2* by a tyrosine kinase raises questions about the role of *cdc2* in oncogenic transformation. In yeast, and apparently in higher eukaryotes, *cdc2* regulates passage through the division cycle and there is no evidence that is involved in growth control. Although *cdc2* has not been isolated as an oncogene, it is perhaps not surprising that the product is a tyrosine kinase substrate, as one element in oncogenesis must be the activation of cell-cycle pathways.

It is generally assumed that mammalian cells monitor the extracellular environment only in G1 and after the initiation of DNA synthesis the cell is fully committed to completion of the division cycle. Thus, it might seem curious that tyrosine kinases,

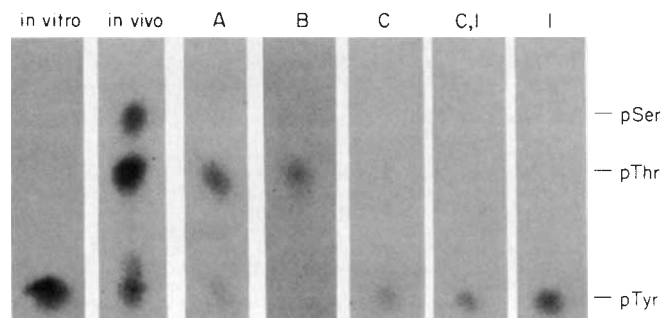


Fig. 7 One-dimensional phosphoamino-acid analysis of p34 labelled *in vivo*, or *in vitro* by pp60^{c-src}. The left-hand two panels represent analysis of the whole protein, whereas those on the right are derived from individual phospho-peptides eluted from two-dimensional chromatographs of tryptic digests. The numbers and letters correspond to those in Fig. 6.

Methods. p34 was labelled *in vivo* or *in vitro* by pp60^{c-src} and subjected to phosphoamino-acid analysis. Alternatively, p34 labelled *in vitro* or *in vivo* was digested with trypsin and the resulting phosphopeptides were separated as described in Fig. 6. Individual phosphopeptides designated A, B, C and I in Fig. 6 were eluted from the cellulose by water, lyophilized and then subjected to phosphoamino-acid analysis. One-dimensional phosphoamino-acid analysis was performed essentially as described⁴³.

presumably involved in signalling growth regulatory information, act on an effector of the G2/M transition. There are three important points in this context. First, even if the initiation of mitosis were an event which was insensitive to environmental cues, it involves drastic rearrangement of cellular structures. This process is undoubtedly highly regulated by intracellular signals and a tyrosine kinase might be involved. Second, the studies that established the pervasive view of the exclusive role of G1 control in mammalian cells were generally undertaken with crude serum rather than purified growth factors. As these factors have been purified and individually characterized, both G1 and G2 effects have been uncovered. For example, B lymphocytes, a classic instance of a 'G1 control' cell type³², require separate factors for metabolic activation, G1 progression and cell division³³. Other cells also display regulation of the cell cycle in G2^{34,35} and the importance of G2 control in lower eukaryotes has long been recognized³⁶, even if relatively few examples have been described in mammalian cells³⁷. Finally, *cdc2* is required in fission yeast for both the G1/S and G2/M transitions¹⁴. As the G1 and G2 functions of *cdc2* have been genetically dissociated in this yeast³⁸ (the cyclin-encoding *cdc13* gene interacts with *cdc2* in mitotic control but is dispensable for DNA replication), it may be presumed that the role of *cdc2* at each stage of the division cycle is not precisely identical. Thus, even though vertebrate *cdc2* is a component of the M-phase promoting factor, it is still possible that the protein kinase may play a role in G1 control in multi-cellular organisms.

We gratefully acknowledge the technical assistance of Monika Stein, Linda Rodgers, Jeannie Wiggins and Regina Whitaker. Jean Wang is particularly thanked for providing serum prepared against v-abl and David Pallas for advice on phosphoamino-acid analysis. Ed Harlow, Leonardo Brizuela and Bob Boehr are thanked for critical comments on the manuscript and the artistic assistance of Jim Duffy and Dave Greene is also acknowledged. This work was supported by grants from the NIH.

Received 6 September; accepted 14 November 1988.

- Greenberg, M. & Ziff, E. *Nature* **311**, 433-438 (1984).
- Bravo, R. & McDonald-Bravo, H. *EMBO J.* **4**, 655-661 (1985).
- Bishop, J. M. *Cell* **42**, 23-38 (1985).
- Hartwell, L. H. *Bact. Rev.* **38**, 164-198 (1974).
- Reed, S. I. *Genetics* **95**, 561-577 (1980).
- Lorincz, A. T. & Reed, S. I. *Nature* **307**, 183-185 (1984).

7. Beach, D. H., Durkacz, B. & Nurse, P. *Nature* **300**, 706-709 (1982).
8. Simanis, V. & Nurse, P. *Cell* **45**, 261-268 (1986).
9. Brizuela, L., Draetta, G. & Beach, D. *EMBO J.* **6**, 3507-3514 (1987).
10. Draetta, G., Brizuela, L., Potashkin, J. & Beach, D. *Cell* **50**, 319-325 (1987).
11. Lee, M. & Nurse, P. *Nature* **327**, 31-35 (1987).
12. Hindley, J., Phear, G. A., Stein, M. & Beach, D. *Molec. cell. Biol.* **7**, 504-511 (1987).
13. Draetta, G. & Beach, D. *Cell* **54**, 17-26 (1988).
14. Nurse, P. & Bissett, Y. *Nature* **292**, 558-560 (1981).
15. Dunphy, W., Brizuela, L., Beach, D. & Newport, J. *Cell* **54**, 423-431 (1988).
16. Gautier, J., Norbury, C., Lohka, M., Nurse, P. & Maller, J. *Cell* **54**, 433-439 (1988).
17. Smith, L. D. & Ecker, R. E. *Dev. Biol.* **25**, 232-247 (1971).
18. Masui, Y. & Markert, C. L. *J. exp. Zool.* **177**, 129-146 (1971).
19. Labbe, J., Lee, M., Nurse, P., Picard, A. & Doree, M. *Nature* **335**, 251-254 (1988).
20. Arion, D., Meijer, L., Brizuela, L. & Beach, D. *Cell* **55**, 371-378 (1988).
21. Draetta, G. (in preparation).
22. Evans, T., Rosenthal, E. T., Youngman, J., Distel, D. & Hunt, T. *Cell* **33**, 389-396 (1983).
23. Morrison, D. K., Browning, P. J., White, M. F. & Roberts, T. M. *Molec. cell. Biol.* **8**, 176-185 (1988).
24. Chackalaparampil, I. & Shalloway, D. *Cell* **52**, 801-810 (1988).
25. Jore, R., Kornbluth, S. & Hanafusa, H. *Cell* **50**, 937-943 (1987).
26. Lischwe, M. A. & Ochs, D. *Analyt. Biochem.* **127**, 453-457 (1982).
27. Hindley, J. & Phear, G. A. *Gene* **31**, 129-134 (1984).
28. Miake-Lye, R. & Kirschner, M. *Cell* **1**, 165-175 (1985).
29. Sunkara, P. S., Wright, D. A. & Rao, P. N. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2799-2802 (1979).
30. Coussens, P. M., Cooper, J. A., Hunter, T. & Shalloway, D. *Molec. cell. Biol.* **5**, 2753-2763 (1985).
31. Spivak, J., Erikson, R. & Maller, L. J. *Molec. cell. Biol.* **4**, 1631-1634 (1984).
32. Melchers, F. & Andersson, J. *Cell* **37**, 715-720 (1984).
33. Melchers, F. & Lernhardt, W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7681-7685 (1985).
34. Gelfant, S. *Symp. Int. Soc. cell. Biol.* **2**, 229-259 (1963).
35. Costello, G., Rodgers, L. & Beach, D. *Curr. Genet.* **11**, 119-125 (1986).
36. Prescott, D. M. *Reproduction of Eukaryotic Cells* 177 (Academic, New York, 1976).
37. Drewinko, B., Yang, L. Y., Barlogie, B. & Trujillo, J. M. *Cell Tissue Kinet.* **17**, 453-463.
38. Booher, R. & Beach, D. *EMBO J.* **6**, 3441-3447 (1987).
39. Beemon, K. & Hunter, T. *J. Virol.* **28**, 551-566 (1978).
40. Piwnicka-Worms, H., Kaplan, D. R., Whitman, M. & Roberts, T. M. *Molec. cell. Biol.* **6**, 2033-2040 (1986).
41. Summers, M. D. & Smith, G. E. in *A Manual of Methods for Baculovirus Vectors and Cloned Insect Cell Culture Procedures* (Texas Agricultural Station, College Station, Texas, 1987).
42. Hunter, T. & Sefton, B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311-1315 (1980).
43. Pallas, D. P. & Solomon, F. *Cell* **30**, 407-414 (1982).

Activation and repression of transcription by homoeodomain-containing proteins that bind a common site

James B. Jaynes & Patrick H. O'Farrell

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

The product of the fushi tarazu (ftz) gene is shown to be a site-dependent activator of transcription. In vitro-defined binding sites act as ftz-dependent enhancers in cultured cells. Another homoeodomain-containing protein, the engrailed gene product, competes for homoeodomain-binding sites and counteracts ftz activation.

GENETIC analysis in *Drosophila melanogaster* has identified a group of regulatory genes that participate in embryonic pattern formation^{1,2}. Many of these genes encode a 60 amino-acid sequence, the homoeodomain, that is highly conserved through evolution³⁻⁵. This domain contains a putative helix-turn-helix motif⁶ such as that found in a number of bacterial regulators⁷, and possesses a sequence-specific DNA-binding activity^{8,9}. Numerous regulatory interactions occur among homoeodomain-encoding genes during development. For example, a combination of segmentation and homoeotic gene products, many of which contain homoeodomains¹, specify the developmental fates of cells in striped patterns^{10,11}. Furthermore, homoeodomain-containing proteins (homoeodomain proteins) regulate both each other¹² and downstream ('cytodifferentiation'¹³) genes. Consequently, it has been attractive to think of homoeodomain proteins as direct regulators of transcription. Due to the complexity of cross-regulatory interactions however, mutations altering or eliminating one homoeodomain protein generally change the expression patterns of many other regulators, confounding efforts to distinguish direct from indirect regulation. Additionally, suspected target genes contain large *cis*-acting control regions that respond, directly or indirectly, to complex combinations of regulators¹⁴⁻¹⁷. As a result, identification of target genes directly regulated by homoeodomain proteins has been problematic.

Previous studies have documented instances in which homoeodomain proteins control transcription. The enhancer activity of a DNA fragment from the *ftz* gene, a homoeodomain-containing member of the pair-rule class of segmentation genes, suggested that the *ftz* gene product (*ftz*) activates its own expression in the embryo¹⁸. In a different approach a *Drosophila* tissue culture system was used to show that the *Ultrabithorax* (*Ubx*) gene product induces expression of its own promoter, and inhibits the *Antennapedia* (*Antp*) (P1) promoter

(M. Krasnow and D. S. Hogness, personal communication, see ref. 19 for summary). In these cases however, the complexity of the regulatory circuitry and of the responding control regions leaves open the possibility that the homoeodomain proteins produce these transcriptional effects through intermediaries. To circumvent the complexities of natural target genes, we chose to use binding sites identified *in vitro* to build homoeodomain-responsive promoters. We find that *ftz* is a potent activator of transcription from these promoters in cultured cells.

During sequential subdivision of the *Drosophila* embryo to produce segmentally repeating patterns, combinations of regulatory gene products direct patterned expression of other regulators. How might these multiple inputs be integrated? At least where regulatory interactions between homoeodomain proteins are involved, hints at one important consideration have already been obtained. *In vitro* binding studies have shown that several distinct homoeodomains that have diverged considerably in their amino-acid sequences bind to the same DNA elements (homoeodomain-binding sites), albeit with somewhat different affinities^{9,20}. This suggests that homoeodomain proteins with similar binding specificities may function by interacting with a family of related sites. If these proteins compete for binding sites and have distinct regulatory consequences once bound, target genes could respond differently to different combinations and/or different levels of the regulators.

Simple target promoters built from synthetic binding sites allow us to show that homoeodomain proteins can compete functionally in cultured cells. In contrast to *ftz*, expression of the *engrailed* (*en*) gene product (*en*) alone has no apparent effect on these promoters. However, co-expression of *en* with *ftz* counteracts *ftz* activation. Thus the regulatory effect of a homoeodomain protein can be influenced by a second homoeodomain protein.

Fig. 1 Plasmids. *a*, Homoeodomain protein producer plasmids (Ac-ftz and pAc-en) contain complete cDNA coding sequences inserted into Marc Krasnow's actin 5C promoter/polyadenylation signal vector, pP_{Ac}. Homoeodomains are the solid black regions. Ac-ftz (pP_{Ac}G1100) was provided by Gary Winslow and Matthew Scott. Stop codons were introduced into the coding sequences to produce pAc-ftz^{STOP} and pAc-en^{STOP} (see below).

b, Responder plasmids 1 and 2 containing truncated distal ADH promoters⁴⁶ (from -33 and -86 nucleotides, respectively, to +53 nt; -86dADH includes a binding site for a factor that appears to be important in regulating ADH expression in the embryo⁴⁷) driving chloramphenicol acetyltransferase (CAT) gene expression (pD-33CAT, pD-86CAT) were provided by Bruce England and Robert Tjian. Six tandem copies of a consensus homoeodomain-binding site⁹ (NP6, see *c* below) were inserted immediately upstream of these promoters. Responders 3 and 4 contain the hsp70 promoter⁴⁸ driving β -galactosidase (β -gal) expression; the latter includes *cis*-acting heat-shock transcription factor binding sites^{49,50}. Various homoeodomain-binding sites were inserted just upstream by Jean-Paul Vincent (J.-P. Vincent and P.H.O'F., unpublished results). Responder 5 (pAF0) contains a proximal ADH promoter and structural gene⁵¹. Again, NP6 was inserted at the 5'-end of the promoter. *c*, Sequences of homoeodomain-binding sites⁹ used in the responder constructs are shown. Tandem arrows show the relative orientations of the individual NP sites within the multimers. No orientation can be assigned to the palindromic RP and LP sites. *d*, Reference plasmids. pCOPiCAT contains a copia transposable element long terminal repeat (LTR) (including the copia promoter)²¹ fused to a CAT structural gene. hsp82LacZ (pLac82SU) was provided by Dale Dorsett (Sloan-Kettering Cancer Center, New York) and contains the hsp82 promoter driving a β -gal structural gene.

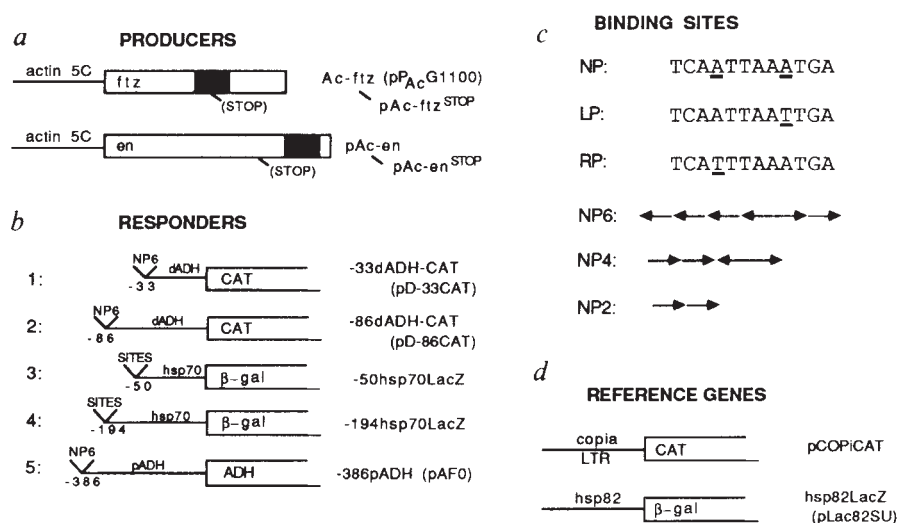
Methods. pAc-en was made by inserting a blunt-ended *Eco*RI fragment, containing *en* cDNA leader, complete coding, and 180 nucleotides of 3' non-translated sequences⁵², into the unique *Bam*HI site of pP_{Ac}. An *Xba*I nonsense codon linker (NEB) was inserted into the unique *Mlu*I site following *en* amino-acid 406 (out of 552) to yield pAc-en^{STOP}. A second nonsense insertion after *en* amino-acid 81 (*Not*I site) was also constructed and tested as a negative control (see text). The *en* homoeodomain (amino acids 453-512) should not be produced by these 'STOP' plasmids. pAc-ftz^{STOP} was similarly constructed from Ac-ftz by insertion of a nonsense codon linker into the unique *Xho*I site after the 16th amino acid of the homoeodomain⁶. A second nonsense insertion after the 33rd amino acid of the homoeodomain (*Eco*RV site) was also made and tested for activity (see text). Both of these 'STOP' constructs are expected to produce a ftz protein truncated before the putative helix-turn-helix DNA-binding motif. To make the '+ sites' versions of responders 1 and 2, a *Sma*I-*Pst*I fragment from an M13mp18 clone containing NP6 in the *Bam*HI site⁹ was inserted into pD-33/-86CAT cut with *Xba*I (blunted) and *Pst*I. Construction of responders 3 and 4 will be described in detail elsewhere. Briefly, they consist of a P-element vector (pSXhLac-7, obtained from Pieter Wensink, Brandeis University) with hsp70 promoter and leader sequences⁴⁸ (-194 to +1,250 nucleotides) fused to the *Escherichia coli* LacZ structural gene, followed by hsp70 3' non-coding sequences. Homoeodomain-binding sites were inserted into unique restriction sites immediately upstream of the promoter. Responder 5 contains the *Drosophila melanogaster* proximal ADH promoter and structural gene starting at -386 nucleotides⁵¹. A *Pst*I-*Eco*RI (blunted) fragment from the M13mp18 clone containing NP6 was inserted into the unique *Pst*I site immediately upstream of the proximal ADH promoter. pCOPiCAT was constructed by fusing a bluescript vector (Stratagene; cut with *Eco*RI and *Bam*HI) with both the copia LTR (*Eco*RI-*Hind*III fragment) from pCOPneo²¹ and the CAT gene/poly(A)-site region (*Hind*III-*Bam*HI fragment) from pSV₂CAT⁵³.

A ftz-dependent enhancer

We inserted homoeodomain-binding sites⁹ adjacent to characterized promoters (Fig. 1) in an attempt to make these promoters responsive to homoeodomain proteins. These constructs (responders) were co-transfected into cultured *Drosophila* cells (Schneider line 2)²¹ either alone or with constructs producing homoeodomain protein (producers). The activity of a responder was assessed by the level of an enzyme activity, the expression of which was driven by the responding promoter.

A responder, consisting of six homoeodomain-binding sites (NP6) upstream of a truncated promoter (the distal alcohol dehydrogenase promoter with 33 base pairs (bp) of upstream sequence: -33dADH), showed a 300-fold increase in activity when co-transfected with the ftz producer (Fig. 2). This increase in expression was entirely dependent on the ftz open reading frame in the producer, as shown by the insertion of stop codons at two different positions in the coding region (see legends of Figs 1 and 2). In addition, induction of transcription was strongly dependent on the homoeodomain-binding sites in the responder (Fig. 2). We conclude that activation by ftz is mediated via the inserted sites. As the inserted sequences are sites for *in vitro* binding by ftz, we suggest that the activation involves direct binding of ftz to the *cis*-acting sites in the responding promoter.

To test the generality of this effect, we constructed responsive promoters by placing the same DNA fragment containing the

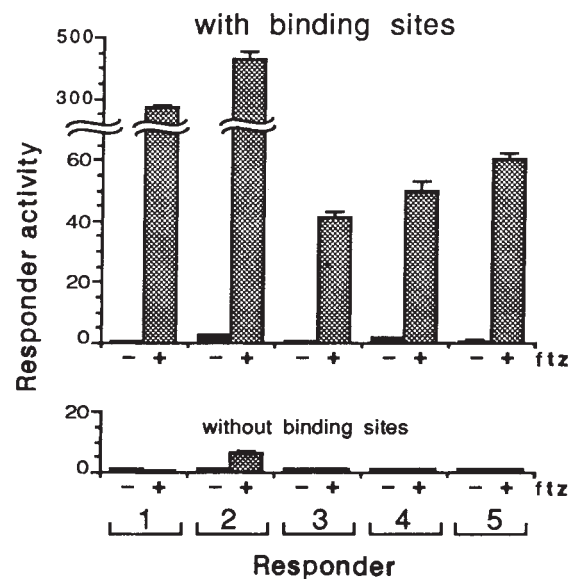


homoeodomain-binding sites (NP6) in other positions and adjacent to other promoters (Fig. 1*b*). In addition to the -33dADH promoter, we used a longer version, -86dADH. The -86dADH promoter had a higher basal level of expression than -33dADH (about 10-fold), but when homoeodomain-binding sites were included immediately upstream, it was also induced by ftz several hundredfold above its basal level (responder 2, Fig. 2). In addition, two configurations of the heat-shock 70 (hsp70) promoter were tested, one with sequences to -50 nucleotides, and one to -194 nucleotides. When homoeodomain-binding sites were added (at -50 or -194), ftz caused a 30- to 60-fold induction (responders 3 and 4, Fig. 2). In the -50 nucleotide configuration, the binding-site array was in the opposite orientation to that in the other responders. The fifth responder, the proximal ADH promoter to -386 nucleotides, gave a 60-fold binding-site-dependent ftz induction (Fig. 2). The basal activity levels of the responders (without ftz) were not significantly affected by the addition of homoeodomain-binding sites.

Although a strong ftz activation (30- to 500-fold) always depended on the presence of homoeodomain-binding sites, much smaller effects of ftz were sometimes seen in the absence of sites. In particular the -86dADH promoter (responder 2, Fig. 2) gave a ftz-dependent stimulation of about seven-fold. Some other promoters also exhibited small but significant stimulation when co-transfected with higher amounts of the ftz pro-

Fig. 2 Ftz induces transcription in cultured cells. *Drosophila* cells (Schneider line 2, cultured as previously described⁵⁴) were co-transfected with (1) either (+ftz) 'producer' plasmid expressing ftz cDNA, or (-ftz) a control plasmid (either pAc-ftz^{STOP} or pP_{Ac}), and (2) the indicated 'responder' plasmid (#1-#5, see Fig. 1), either with (upper panel) or without (lower panel) the NP6 homoeodomain-binding sites. Responder activity was ascertained by quantitating the appropriate enzymatic activity in whole-cell extracts. Activity was determined relative to the activity of a co-transfected reference gene plasmid; the results were essentially the same when expressed per μ g of total protein in the extract (determined by the method described⁵⁵). Activity is normalized to the basal level of each responder, without binding sites and '-ftz' (set at 1.0). Error bars indicate the range of values in duplicate transfections within the same experiment. Equivalent results were obtained in at least two separate experiments for each responder.

Methods. Cells were grown in Schneider's *Drosophila* medium⁵⁴ supplemented with 12% fetal bovine serum. Transfection was by the calcium phosphate precipitation method, done essentially as described⁵⁴. Cells were collected two days after transfection by scraping, rinsed once with phosphate-buffered saline and lysed by three cycles of freeze-thaw (-70°C to room temperature). Extracts were prepared by centrifugation (15,000g, 5 min), and supernatants were either stored at -70°C or assayed immediately. CAT activity (from responders 1 and 2, and pCOPICAT reference gene used with responders 3 and 4) was determined as described⁵⁶. Relative activities were the same whether or not extracts were heated (68°C , 5 min) to destroy potential de-acetylase activity⁵⁷. High activity extracts were diluted to keep determinations within the linear range of the assay, as shown by standard curves. CAT reaction time courses and extract mixing experiments indicated that little or no inhibitory activity was present in low activity extracts (data not shown). β -galactosidase (β -gal) activity (from responders 3 and 4 and hsp82LacZ reference gene used with responders 1, 2 and 5) was determined by fluorimetric assay⁵⁸. *Drosophila* ADH activity was determined as described⁵⁹. There was a detectable background activity in cells from control transfections (without responder plasmid) for β -gal and ADH (but not CAT). The β -gal background was at most 10% of the smallest plasmid-dependent activity, and so did not affect the results presented. The ADH background was relatively higher, about equal to the uninduced activities shown (responder 5), so the induction value for the -386pADH promoter (60-fold) is a minimum estimate. Amounts of plasmid DNA used per transfection (60-mm Petri dish) were as follows: 5 μ g responder 1 (+/- sites) with 1 μ g Ac-ftz (+/- STOP codon), 1 μ g hsp82LacZ (reference gene), and 3 μ g pAc-ftz^{STOP}; 1 μ g responder 2 (+/- sites) with either 1 μ g Ac-ftz and 8 μ g pP_{Ac}, or 9 μ g pP_{Ac}; 0.1 μ g responder 3 (+/- sites) with 3 μ g Ac-ftz (+/-STOP codon), 1 μ g pCOPICAT (reference gene), and 3 μ g pP_{Ac}; 0.05 μ g responder 4 (+/- sites) with 3 μ g Ac-ftz (+/- STOP codon), 1 μ g pCOPICAT (reference gene), and 6 μ g pAc-ftz^{STOP}; or, 7 μ g responder 5 (+/- sites) with 3 μ g Ac-ftz (+/-STOP codon), and 0.01 μ g hsp82LacZ (reference gene). Amounts of responder DNA were chosen to give a readily quantifiable basal (uninduced) activity level, except responder 5 (see above). Small amounts of responders 3 and 4 were used because much larger amounts (1-2 μ g) gave somewhat reduced induction ratios, consistent with either a titration of activating factor(s) or a 'ceiling' on stable β -gal accumulation in the cells.



ducer (not shown). These effects could be due to adventitious ftz-binding sites in these promoters, or there might be small site-independent effects on some promoters at relatively high ftz concentrations. Nonetheless, the major effects of ftz were clearly site dependent.

We conclude that the homoeodomain-binding sites act as ftz-dependent enhancers²². That is, they act on all promoters tested, in either orientation, and at distances up to at least 386 nucleotides, to activate transcription.

Enhancer strength

Enhancer activity generally depends on the presence of multiple discrete elements which act synergistically²³⁻²⁷. The ftz-dependent enhancer used in the previous experiments consists of six tandem copies of a 12-nucleotide sequence. To test the importance of multiple copies we examined ftz activation of the -50hsp70 promoter with either six, four or two consensus copies (NP6, NP4, NP2, Fig. 3a). The reduction from six copies to four resulted in a large decrease in ftz-dependent activation, from about 60-fold to about 10-fold. With only two consensus copies, activation was reduced further to about two-fold. A similar response to the number of NP copies was seen with the -194hsp70 promoter (data not shown).

The magnitude of the effect of reducing binding-site copy number from six to four was somewhat unexpected. Although *in vitro* studies suggested cooperative binding, the major effect was seen between one and two copies. An increase from four to six copies only marginally improved *in vitro* binding⁹. Perhaps site occupancy *in vivo* has a different dependence on the number of copies. Alternatively, the dramatic difference between transcriptional activation by four versus six copies might not represent a correspondingly large difference in site occupancy. Rather, an increase in the number of ftz-occupied sites might produce a disproportionately large increase in transcription. In

either case, the copy number effect seen here is reminiscent of the cooperative or synergistic action of multiple enhancer elements in other systems.

If ftz activation requires interaction with specific binding sites in a responding gene, then site alterations that change binding affinity should also produce an altered ftz response. We therefore tested the effectiveness of two homoeodomain-binding sites, each altered from the consensus by a single nucleotide. These changes create a palindromic sequence related either to the left half (LP) or the right half (RP) of the original consensus (NP, see Fig. 1). Tandem repeats of four copies of each of these altered sequences (LP4, RP4) were compared to four copies of the original consensus (NP4). *In vitro* studies had shown that the change from NP to LP had no detectable effect on ftz-binding affinity, whereas the NP to RP change caused about a five-fold reduction in affinity⁹. Consistent with the *in vitro* results, LP4 and NP4 showed about the same ability to confer ftz activation to the -50hsp70 promoter, whereas RP4 had significantly reduced activity (Fig. 3b). Similar results were obtained with the same binding sites upstream of the -194hsp70 promoter (for example, Fig. 4b, data in the absence of en, 0 en). Thus these binding site alterations have parallel effects on *in vitro* binding affinity and transcriptional activation. This parallel behaviour supports our proposal that ftz binding is a prerequisite for activation.

Engrailed represses ftz activation

The homoeodomain-binding site used to demonstrate ftz-dependent activation is also bound by the en protein *in vitro*^{8,9}. Nonetheless, when en producer plasmid was co-transfected with the responder constructs that are inducible by ftz, no activation occurred. In some cases, a small but reproducible reduction in expression (two to threefold) was observed with, but not without, binding sites in the responders (data not shown).

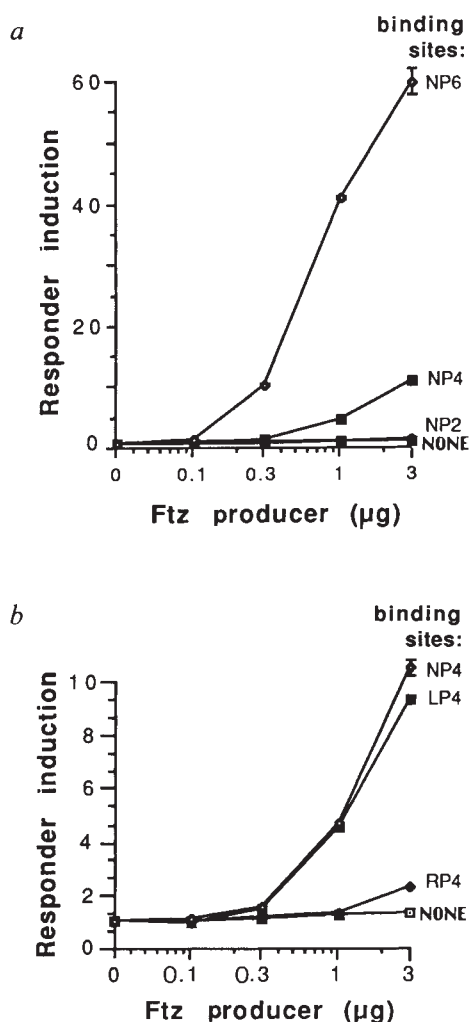


Fig. 3 Alterations in binding sites change responsiveness to ftz. Different amounts of ftz producer (see below) were co-transfected with the -50hsp70 responder containing, *a*, either zero, two, four, or six NP binding sites (see Fig. 1), and, in *b*, four tandem copies of either the consensus (NP), or single-nucleotide alterations (LP, RP; see Fig. 1), or no homoeodomain-binding sites. Activity of each responder was determined as in Fig. 2, and was normalized to the basal level without ftz, set at 1.0, to give 'Responder induction'. Error bars indicate the range of values of duplicate transfections. Similar results were obtained with the -194hsp70 responder. Amounts of plasmid DNA used per transfection were: 0.1 µg each responder with 3 µg pP_{Ac}, 1 µg pCOPiCAT (reference gene), and the amounts of ftz producer (Ac-ftz) shown. Total DNA was made constant by adding pAc-ftz^{STOP}. Control experiments using a plasmid with an actin 5C promoter driving CAT showed that expression levels were proportional to the amount of added DNA from 0.1 to 10 µg (data not shown).

To test whether en could bind to the same sites as ftz *in vivo*, we produced ftz and en together, reasoning that if en bound, it might affect ftz activation. Indeed, with each of the binding-site-containing responders, ftz activation was reduced between 30 and 100-fold by the presence of en (Fig. 4a). This effect of the en producer is entirely dependent on the en open reading frame, as insertion of stop codons at either of two different positions in the coding region abolished effectiveness (see Fig. 1 legend and Fig. 4). Thus, the combination of en with ftz can have an effect on transcription that is qualitatively distinct from that of ftz alone.

We examined the degree of en antagonism of ftz induction using different amounts of transfected en producer. Whether responders contained NP6 (data not shown), NP4, or LP4

binding sites, ftz induction was similarly titrated by increasing amounts of en producer (Fig. 4b). This is consistent with the fact that en, like ftz (see above), binds equally well to these sites *in vitro*⁹. As mentioned above, the *in vitro* affinity of ftz for RP4 is less than for NP4 (about fivefold), and ftz activation of RP4-containing responders in transfected cells is less than activation of NP4 responders. The *in vitro* affinity of en for RP4 is also lower than for NP4 (about 25-fold), and although the RP4 responder is induced less than three-fold by ftz, en is still able to significantly antagonize this activation (Fig. 4b). Consistent with the fact that, *in vitro*, the NP to RP site alteration reduces en affinity more than it reduces ftz affinity (25- compared to fivefold), the en repression with RP4 seems to be less effective than with NP4 (that is, Fig. 4b, the RP4 and NP4 curves cross). Although this distinction is reproducible (data not shown), its significance is, nonetheless, uncertain because it relies on comparison of small differences in the activities of the responders when repressed by en.

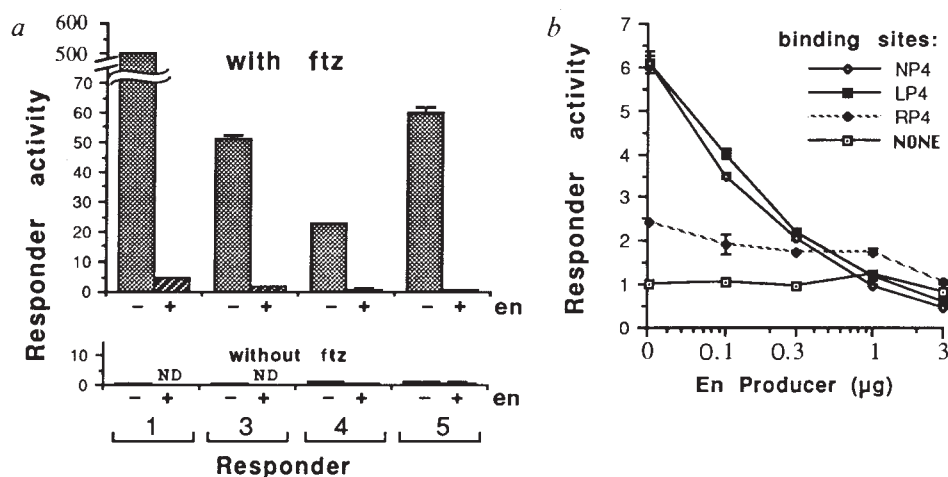
Discussion

The results presented here have shown that ftz is a powerful site-dependent activator of transcription. Sequences identified as *in vitro* binding sites for bacterially produced homoeodomain proteins⁹, when inserted upstream of promoters, act as ftz-dependent enhancers in cultured cells. This suggests that ftz acts by binding to the inserted sites to activate the linked promoter. However, we cannot formally eliminate the possibility that ftz is acting indirectly by altering the activity of another regulator that functions at the sites.

Another homoeodomain protein, en, is also able to bind *in vitro* to the sites that confer ftz responsiveness in our assay⁹. Although we do not observe transcriptional activation by en in cultured cells, we find that en negates activation by ftz. The parallel of *in vitro* binding and functional antagonism *in vivo* suggests the simple hypothesis that binding competition might be responsible for repression by en. That is, en, by binding to the sites, may prevent ftz binding and produce a complex that is ineffective at activating transcription (or possibly a complex that actively represses transcription). There are however, numerous alternative explanations for the observed repression. Perhaps en interacts with and inactivates ftz, or perhaps en interferes with accessory or intermediary factors necessary for ftz action.

Although we cannot yet specify mechanisms for the transcriptional activation by ftz or for the functional antagonism by en, we favour the simple model that both proteins act directly at the homoeodomain-binding sites. If such a model is correct, we might expect that other homoeodomain proteins having related DNA-binding specificities *in vitro* would affect transcription from the same promoters that are activated by ftz. Indeed, in preliminary experiments the Ubx gene product activated expression in a binding site (NP6)-dependent manner (albeit less effectively than ftz). Outside the homoeodomain, which specifies DNA-binding activity, Ubx and ftz share no significant homology. As Ubx can bind to homoeodomain-binding sites also bound by ftz (Gavis and D. S. Hogness, personal communication), this result is compatible with direct action of these proteins on the target sequences. In addition, we have found that transfection with a combination of ftz and Ubx producers activated the responder to levels intermediate to those obtained with ftz or Ubx alone. Such an averaging effect could occur if the proteins were similar in that they compete for the *cis*-acting sites, but different in that they produce complexes activating transcription to different degrees. A fourth homoeodomain protein, the *even-skipped* (eve) gene product, which has a binding specificity related to that of ftz, was also found to alter transcription in a binding site (NP6)-dependent fashion. Eve, like en, was inactive on its own but strongly antagonized induction by ftz and Ubx. As several members of the homoeodomain family of regulators seem to act on the same sites in our transfection

Fig. 4 Repression by en. Error bars indicate the range of values from duplicate transfections in the same experiment. Equivalent results were obtained in at least two separate experiments. *a*, Cells were co-transfected (see Fig. 2 legend) with (1) producer plasmids Ac-ftz ('with ftz'), pAc-en (+en), pAc-ftz^{STOP} ('without ftz'), and/or pAc-en^{STOP} (-en), and (2) the indicated responder (see Fig. 1) with the NP6 homoeodomain-binding sites. Without homoeodomain-binding sites, there was no response to either ftz (Fig. 2) or en (data not shown). Amounts of plasmid DNA used per transfection were as follows: 5 μ g responder 1 with 1 μ g hsp82LacZ (reference gene) and either (1) (-ftz, -en) 4 μ g pAc-ftz^{STOP}, (2) (+ftz, -en) 1 μ g Ac-ftz and 3 μ g pAc-en^{STOP}, or (3) (+ftz, +en) 1 μ g Ac-ftz, 1 μ g pAc-en, and 2 μ g pAc-en^{STOP}; 0.1 μ g responder 3 with 1 μ g pCOPiCAT (reference gene) and either (1) (-ftz, -en) 3 μ g pAc-ftz^{STOP} and 3 μ g pAc-en, (2) (+ftz, -en) 3 μ g Ac-ftz and 3 μ g pAc-en^{STOP}, or (3) (+ftz, +en) 3 μ g Ac-ftz and 3 μ g pAc-en; 0.05 μ g responder 4 with 0.2 μ g pCOPiCAT (reference gene), 3 μ g Ac-ftz (+/-STOP codon, = 'without/with ftz'), and 6 μ g pAc-en (+/-STOP codon, = '-/+en'); 7 μ g responder 5 with 0.01 μ g hsp82LacZ (reference gene) and either (1) (-ftz, -en) 3 μ g pAc-ftz^{STOP}, (2) (-ftz, +en) 2 μ g pAc-ftz^{STOP} and 1 μ g pAc-en, (3) (+ftz, -en) 1 μ g Ac-ftz, 1 μ g pAc-ftz^{STOP}, and 1 μ g pAc-en^{STOP}, or (4) (+ftz, +en) 1 μ g Ac-ftz, 1 μ g pAc-ftz^{STOP}, and 1 μ g pAc-en. 'ND', response with en alone not determined in this experiment; several other experiments with responder 1 showed that en had a small negative effect with binding sites (see text). *b*, Different amounts of en producer were co-transfected with a constant amount of ftz producer and the -194hsp70 responder containing either no homoeodomain-binding sites, or the NP4, LP4, or RP4 binding site array (see Fig. 1). Amounts of plasmid DNA per transfection were: 0.05 μ g each responder with 3 μ g Ac-ftz, 1 μ g pCOPiCAT (reference gene), and the amounts of en producer (pAc-en) shown. Total DNA was made constant with pAc-en^{STOP}. Responder activities for both *a* and *b* were determined as described in the Fig. 2 legend.



assay, we propose that they act by directly binding to the sites, and that when present together they compete for binding sites and influence transcription according to the activity of the resulting complex.

Several precedents exist for competition among regulators for *cis*-acting sites. For example, in the lysis/lysogeny switch of bacteriophage λ , cro and repressor compete for related sites in the rightward operator²⁸. In eukaryotes, a number of regulatory proteins have been shown to be members of protein families with related DNA-binding specificities²⁹⁻³⁷. In fact, a member of the nuclear hormone receptor family, thyroid hormone receptor, was recently shown to mediate negative effects on transcription from oestrogen-responsive DNA elements, apparently by competing with oestrogen receptor for binding to the elements³⁸. It is possible that regulatory gene families in general evolve by duplication and divergence to produce members with overlapping functions. The modular nature of protein structure (for example, see ref. 39) would make it likely that some domains retain their original function whereas others diverge by losing one interaction and possibly gaining a new functional interaction. The members of such protein families would interdependently regulate target genes because of an overlapping function (such as DNA binding to the same family of sites), but would have differing effects based on a diverged function (such as activation versus repression of transcription).

Competition for binding sites could dramatically influence the action of homoeodomain proteins during embryonic pattern formation. Each of the many known homoeodomain proteins is expressed in a complex pattern during embryogenesis. As a result of overlap of these patterns, binding sites are exposed to a mixture of homoeodomain proteins that varies in composition in a complex spatio-temporal fashion. As the outcome of binding competition will depend on the relative concentrations of the binding proteins, the enhancer activity of a site might also vary in a detailed spatio-temporal pattern. Additionally, the outcome of competition depends on the relative affinity of each particular site for the different binding proteins. As a result, each of the related binding sites might undergo transitions in occupancy at distinct relative concentrations of homoeodomain proteins and would consequently have a unique spatio-temporal programme of activity (Fig. 5).

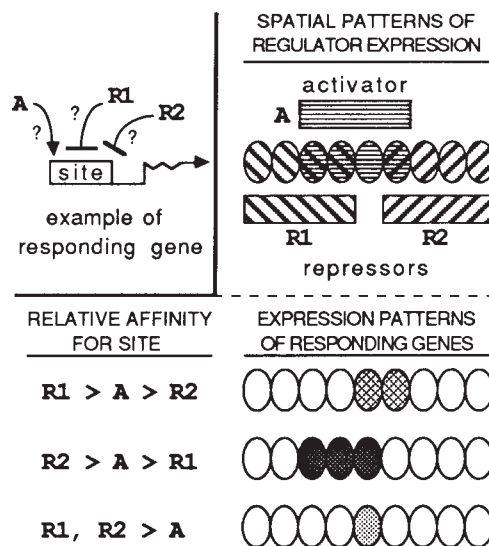


Fig. 5 Multiple patterns of gene expression can be generated by competition among regulators for binding to related but different *cis*-acting sites. Beginning with overlapping patterns of expression of three regulatory proteins with related DNA-binding specificities, three novel patterns are produced by differential competition. A is an activator, whereas R1 and R2 are repressors. Related *cis*-acting sites in the responding genes have different relative affinities for the regulators. The first responding gene has a binding site with relative affinities R1 > A > R2, so that it is activated by A and repressed by R1, but not by R2. The second responding gene has a site with relative affinities R2 > A > R1, whereas the site in the third gene has a higher affinity for both R1 and R2 than for A. Such a scheme could operate in the embryo to refine spatial patterns of regulatory gene expression and to subdivide domains of expression. This simple example assumes that regulator concentrations are equal and constant in space and time. Similar competitive interactions could also generate multiple domains of responder expression from a gradient of regulator concentration.

Although the potential exists in the embryo for competition to occur among homeodomain proteins for binding sites, the situation is likely to be more complicated than that in cultured cells. For example, other work performed in our laboratory has shown that promoters responsive to *ftz* in culture do not produce a *ftz*-dependent pattern of expression in the embryo (J.-P. Vincent, personal communication). A number of interactions or modifications might occur during embryogenesis to change the behaviour of regulators. For example, phosphorylation can potentially produce⁴⁰ or unmask^{41,42} acidic domains that activate transcription⁴³. Accordingly, *ftz* and *en*, both of which are known to be phosphorylated^{44,45}, might not always have the same transcriptional effects that we have observed in these experiments.

As yet, we have no direct evidence that competitive interactions among transcription factors are important in embryonic

development. Nonetheless, the prevalence of a highly conserved DNA-binding domain among developmental regulators leads us to propose that competitive interactions will have a general role in control mechanisms governing spatial- and tissue-specific transcription.

We thank Jill Jongens, John Little, Delia Lakich, Rich Kostriken, Jean-Paul Vincent, Bruce Edgar, Bill Kalionis and Steve DiNardo for critical advice on the manuscript. Additionally, we thank Claude Desplan and all members of the O'Farrell laboratory for help and encouragement. We thank Jean-Paul Vincent, Gary Winslow, Mark Krasnow, Anne Ephrussi and Dale Dorsett for sharing plasmid constructs and information before publication. We would also like to acknowledge our debt to Mark Krasnow whose earlier work stimulated these studies. This work was funded by an NIH grant to P.H.O. and by an ACS post-doctoral grant to J.B.J.

Received 20 September; accepted 14 November 1988.

1. Scott, M. P. & O'Farrell, P. H. *A. Rev. cell. Biol.* **2**, 49-80 (1986).
2. Scott, M. P. & Carroll, S. B. *Cell* **51**, 689-698 (1987).
3. Scott, M. P. & Weiner, A. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4115-4119 (1984).
4. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. *Cell* **37**, 403-408 (1984).
5. Gehring, W. J. & Hiromi, Y. *A. Rev. Genet.* **20**, 147-173 (1986).
6. Laughon, A. & Scott, M. P. *Nature* **310**, 25-31 (1984).
7. Pabo, C. O. & Sauer, R. T. *A. Rev. Biochem.* **53**, 293-321 (1984).
8. Desplan, C., Theis, J. & O'Farrell, P. H. *Nature* **318**, 630-635 (1985).
9. Desplan, C., Theis, J. & O'Farrell, P. H. *Cell* **54**, 1081-1090 (1988).
10. DiNardo, S. & O'Farrell, P. H. *Genes Dev.* **1**, 1212-1225 (1987).
11. Howard, K. & Ingham, P. *Cell* **44**, 949-957 (1986).
12. Hafen, E., Levine, M. & Gehring, W. *Nature* **307**, 287-298 (1984).
13. Garcia-Bellido, A. *Ciba Fdn Symp.* **29**, 161-182 (1975).
14. Bender, W. *et al. Science* **221**, 23-29 (1983).
15. Kuer, J. M. *et al. Cell* **42**, 309-316 (1985).
16. Beachy, P. A., Helfand, S. L. & Hogness, D. S. *Nature* **313**, 545-551 (1985).
17. Kassiss, J. A., Wong, M. L. & O'Farrell, P. H. *Molec. cell. Biol.* **5**, 3600-3609 (1985).
18. Hiromi, Y. & Gehring, W. J. *Cell* **50**, 963-974 (1987).
19. Robertson, M. *Nature* **327**, 556-557 (1987).
20. Hoey, T. & Levine, M. *Nature* **332**, 858-861 (1988).
21. Rio, D. C. & Rubin, G. M. *Molec. cell. Biol.* **5**, 1833-1838 (1985).
22. Serfling, E., Jasin, M. & Schaffner, W. *Trends Genet.* **1**, 224-230 (1985).
23. DeFranco, D. & Yamamoto, K. R. *Molec. cell. Biol.* **6**, 993-1001 (1986).
24. Herr, W. & Clarke, J. *Cell* **45**, 461-470 (1986).
25. Ptashne, M. *Nature* **322**, 697-701 (1986).
26. Schirm, S., Jiricny, J. & Schaffner, W. *Genes Dev.* **1**, 65-74 (1987).
27. Ondek, B., Gloss, L. & Herr, W. *Nature* **333**, 40-45 (1988).
28. Ptashne, M. *A Genetic Switch—Gene Control and Phage λ* (Cell, Cambridge, Massachusetts 1986).
29. Dorn, A., Bollekens, J., Staub, A., Benoist, C. & Mathis, D. *Cell* **50**, 863-872 (1987).
30. Jones, N. C., Rigby, P. W. J., & Ziff, E. B. *Genes Dev.* **2**, 267-281 (1988).

31. Santoro, C., Mermod, N., Andrews, P. C. & Tjian, R. *Nature* **334**, 218-224 (1988).
32. Von der Ahe, D. *et al. Nature* **313**, 706-709 (1985).
33. Chowdhury, K., Deutsch, U. & Gruss, P. *Cell* **48**, 771-778 (1987).
34. Struhl, K. *Cell* **50**, 841-846 (1987).
35. Lambert, P. F., Spalholz, B. A. & Howley, P. M. *Cell* **50**, 69-78 (1987).
36. Chalepakis, G. *et al. Cell* **53**, 371-382 (1988).
37. Franza, B. R. Jr, Rauscher, F. J., III, Josephs, S. F. & Curran, T. *Science* **239**, 1150-1153 (1988).
38. Glass, C. K., Holloway, J. M., Devary, O. V. & Rosenfeld, M. G. *Cell* **54**, 313-323 (1988).
39. Rusconi, S. & Yamamoto, K. R. *EMBO J.* **6**, 1309-1315 (1987).
40. Sorger, P. K., Lewis, M. J. & Pelham, H. R. B. *Nature* **329**, 81-84 (1987).
41. Yamamoto, K. R. *A. Rev. Genet.* **19**, 209-252 (1985).
42. Sen, R. & Baltimore, D. *Cell* **47**, 921-928 (1986).
43. Giniger, E. & Ptashne, M. *Nature* **330**, 670-672 (1987).
44. Krause, H. M., Klemenz, R. & Gehring, W. J. *Genes Dev.* **2**, 1021-1036 (1988).
45. Gay, N. J., Poole, S. J. & Kornberg, T. B. *Nucleic Acids Res.* **16**, 6637-6647 (1988).
46. Heberlein, U., England, B. & Tjian, R. *Cell* **41**, 965-977 (1985).
47. Heberlein, U. & Tjian, R. *Nature* **331**, 410-415 (1988).
48. Lis, J. T., Simon, J. A. & Sutton, C. A. *Cell* **35**, 403-410 (1983).
49. Pelham, H. R. B. *Cell* **30**, 517-528 (1982).
50. Parker, C. S. & Topol, J. *Cell* **37**, 273-283 (1984).
51. Posakony, J. W., Fischer, J. A. & Maniatis, T. *Cold Spring Harb. Symp. quant. Biol.* **50**, 515-520 (1985).
52. Poole, S. J., Kauvar, L. M., Drees, B. & Kornberg, T. *Cell* **40**, 37-43 (1985).
53. Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044-1051 (1982).
54. Di Nocera, P. P. & Dawid, I. B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7095-7098 (1983).
55. Bradford, M. M. *Analyt. Biochem.* **72**, 248-254 (1976).
56. Kimura, A., Israel, A., Le Bail, O. & Kourilsky, P. *Cell* **44**, 261-272 (1986).
57. Crabb, D. W. & Dixon, J. E. *Analyt. Biochem.* **163**, 88-92 (1987).
58. Stuart, G. W., Searle, P. F., Chen, H. Y., Brinster, R. L. & Palmiter, R. D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7318-7322 (1984).
59. Maroni, G. *Biochem. Genet.* **16**, 509-523 (1978).

LETTERS TO NATURE

Evidence from sub-millimetre observations for thermal dust emission in NGC4151

R. A. Edelson*, W. K. P. Gear†, M. A. Malkan‡ & E. I. Robson§

* Center for Astrophysics and Space Astronomy, Campus Box 391, University of Colorado, Boulder, Colorado 80309-0391, USA

† Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

‡ Astronomy Department, Mathematical Sciences Building, University of California, Los Angeles, California 90024-1526, USA

§ School of Physics and Astronomy, Lancashire Polytechnic University, Preston PR1 2TQ, UK

Most quasars and Seyfert 1 galaxies have spectra that rise from optical wavelengths into the 60–100 μm infrared region, beyond which they must suffer a sharp cutoff indicated by the lack of observed millimetre emission. There is evidence^{1–4} that the infrared emission receives contributions from both thermal emission by interstellar dust and synchrotron radiation from a compact source. Determining the nature of the infrared emission is thus an important step in understanding the structure of quasars and Seyfert 1

galaxies. We report here observations at 438- μm wavelength of the Seyfert 1 galaxy NGC4151 using the UKT14 bolometer on the James Clerk Maxwell Telescope. We find a 5σ upper limit of 200 mJy in an 11" aperture. Comparison with an earlier measurement at 155 μm indicates that at least half of the 155- μm flux is due to thermal dust emission. The remainder may be from a synchrotron source which becomes self-absorbed at wavelengths of less than 80 μm .

A number of indicators have been used to determine the relative importance of the thermal dust and non-thermal synchrotron processes¹, including: (1) infrared spectral shape (dusty objects tend to have steeper spectra); (2) infrared spectral features produced by dust grains (including 10 μm silicate absorption and 3.3 μm emission features); (3) reddening, as measured by, for example, the line ratios, ultraviolet continuum shape and strength of the 2,175 Å absorption feature; (4) infrared variability (only non-thermal sources can vary quickly); (5) infrared polarization (high polarization suggests non-thermal emission); and (6) infrared compactness, as measured by the ratio of the IRAS large-beam 12- μm flux to the ground-based, small-beam 10.6- μm flux (non-thermal sources are compact). However, none of these tests definitively establishes the emission mechanism. Most of the indicators are consistent with, but do not prove, the suggestion that the infrared spectra of the more luminous quasars tend to be dominated by non-thermal

emission, whereas in the lower-luminosity Seyfert 1 galaxies, thermal dust emission can often dominate¹⁻⁴.

Measuring the shape of the far-infrared cutoff can distinguish between thermal and non-thermal emission, because non-thermal synchrotron emission must produce a relatively flat cutoff⁵, with spectral index $\alpha \leq +2.5$ (intensity at frequency ν , $S_\nu \propto \nu^\alpha$) whereas dust emission produces a sharper cutoff⁶, with $\alpha = +3$ to $+4$. A new generation of sub-millimetre telescopes, including the 15-m James Clerk Maxwell Telescope (JCMT), the Kuiper Airborne Observatory (KAO) and the IRAM 30-m telescope, along with new, sensitive bolometers and arrays, are now making it possible to study these sources at sub-millimetre wavelengths and to determine the shape of the cutoff.

NGC4151 (photographic magnitude from Zwicky survey, $m_{\text{pg}} = 11.2$, redshift $z = 0.0030$) is one of the nearest, brightest and best studied Seyfert 1 galaxies. Although it has been detected from γ -ray to radio wavelengths, sensitive observations between 100 μm and 1.5 cm have been made only recently. Engargiola *et al.*⁷ detected NGC4151, at a level of 4.8 ± 0.7 Jy, in the central element of a 155- μm array camera with 55" pixels. They also found evidence for emission in the adjacent pixels, at a level of 5.4 ± 1.4 Jy. Here we report the first observations made at 438 μm of this or any other Seyfert 1 galaxy with the UKT14 bolometer on the JCMT, and we discuss the implications of these results for the nature of the infrared sub-millimetre emission.

The observations were made with the JCMT on 20 April 1988 UT. The detector was the UKT14 system, a single-channel, ³He-cooled, composite-doped germanium bolometer. A band-pass filter was used to define a central wavelength of 438 μm (685 GHz) and a bandwidth of 84 GHz. Maps made of the beam showed that it was approximately gaussian, with a full width at half maximum of 11". Sky subtraction was performed in the usual manner, with the secondary chopping 30" in azimuth at a frequency of 7.8 Hz. Calibration was done relative to Uranus, which was unresolved and had an assumed flux density⁸ of 199 Jy at 438 μm . Repeated observations showed that the atmospheric transmission was $\sim 10\%$. Flux densities were derived assuming a spectral index of $\alpha = +2$, but the relatively narrow bandwidth means that the fluxes of sources with different spectral indices are in error by only a few per cent. Twenty 20-second pairs were combined to form 400-second-long scans, assuming a gaussian distribution of errors.

At the time of these observations, the absolute pointing of the JCMT was only accurate to about 5", so the tracking was continually verified with nearby bright sources. Because NGC4151 lies conveniently between two bright VLA calibrators⁹, 1156+295 and 1216+487, the pointing was updated relative to a different one of them after each 15-min scan. Alternating between these two sources showed that the r.m.s. drift between scans was $\leq 3.7''$ in both coordinates for these calibrators. As the distance from either calibrator to NGC4151 was about half of that between the two calibrators, this is a conservative upper limit on the pointing errors for NGC4151. This indicates that the pointing was accurate to within a few arcseconds throughout the observations, and the loss of signal was apparently less than 30%.

NGC4151 was not detected at 438 μm in these measurements. As this is the first published report of deep 438- μm photometry with JCMT, a 5σ detection limit of 200 mJy is quoted. Figure 1 shows the continuum energy distribution for NGC4151, made without correcting for the different beam sizes used. The limit on the spectral index is found to be $\alpha_{155-438\mu\text{m}} \geq +3.06$, which constrains the shape of the sub-millimetre energy distribution much more strongly than any previous observations. The uncertainty in $\alpha_{155-438\mu\text{m}}$ due to errors in the 155- μm measurement is only 0.13. This places strong limits on the contribution from synchrotron emission, because the sharpest cutoff that could be produced by synchrotron self-absorption would correspond to $\alpha_{155-438\mu\text{m}} = +2.5$.

We must consider the possibility that the source may have

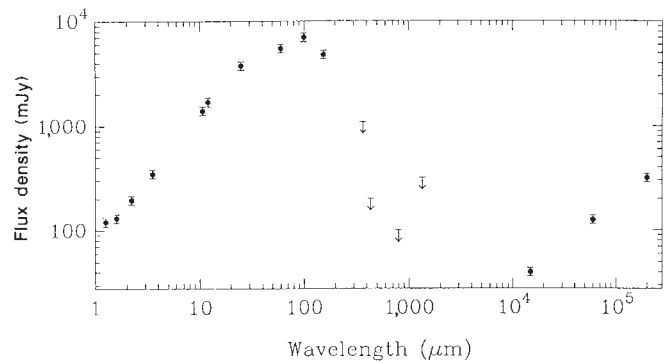


Fig. 1 Infrared-through-radio spectral energy distribution for NGC4151^{1,4,7,14,20}. Flux density is plotted as a function of wavelength (that is, $\log(F_\nu)$ against $\log(\lambda)$). The infrared spectrum is relatively constant, with $\alpha \approx -1.2$ from 1 to 100 μm , and then cuts off sharply at longer wavelengths. The best explanation is that non-thermal synchrotron emission dominates in the near-infrared, but thermal dust emission becomes important at longer wavelengths, dominating in the sub-millimetre. The radio emission appears to be a separate component, with much less power than the infrared.

varied in the year between the 155- μm observations. NGC4151 is known to vary at ultraviolet wavelengths¹⁰, but this appears to be related to a spectral component often called the 'blue bump' (thought to be thermal emission from an accretion disk), which is not important¹¹ for wavelengths $> 1 \mu\text{m}$. Repeated 10- μm observations showed that the flux was constant¹² to within 8% over 1.25 yr in 1978-79, and it did not vary at 12-100 μm , above the 20% level, when observed with IRAS over a one-month period¹³. (In fact, the latter study found no evidence of any far-infrared variations in any Seyfert galaxy or normal quasar for timescales up to six months and longer.) NGC4151 was also constant to better than 10% at radio wavelengths¹⁴ over ~ 5 years. Therefore, it is most unlikely that these sub-millimetre observations were affected by variability.

Another possibility is that the infrared emission could be non-thermal synchrotron emission that is cut off at long wavelengths by free-free absorption in cold intervening material⁵. This would be consistent with the observed sub-millimetre spectrum, because free-free absorption produces an exponential cutoff. The turnover frequency is given by¹⁵ $\nu_t = 0.0003 T_e^{0.75} n_e^{0.5}$ Hz, where T_e is the electron temperature in units of 10^4 K, n_e is the electron density in cm^{-3} and l is the path length in cm. X-ray measurements¹⁶ suggests that a broad-emission-line cloud could lie along the line of sight to NGC4151. The properties of these clouds are not very well constrained by current photo-ionization models. The standard model¹⁷ incorporates an extended partially ionized zone 5×10^{12} cm thick, which, at 8,000 K, is in pressure equilibrium with the fully ionized zone at 20,000 K, and has an average electron density of $\sim 3.5 \times 10^9 \text{ cm}^{-3}$. The consequent free-electron pathlength would produce a free-free absorption turnover $\nu_t \approx 100 \mu\text{m}$. Although such a broad-line cloud along the line of sight could explain the turnover which we detect in NGC4151, it could not explain far-infrared turnovers in many higher luminosity Seyfert 1 nuclei and quasars, which have no detectable soft-X-ray absorption and low broad-line-cloud-covering factors¹⁸ (below 10%). But many of these luminous objects² show turnovers at wavelengths less than 100 μm . It is unlikely that the turnovers in NGC4151 is due to a special mechanism which is different to that responsible for turnovers in most other Seyfert 1 galaxies and quasars.

Other known components with even lower densities, such as the narrow-line region and the postulated hot intercloud medium, could not produce a free-free cutoff at wavelengths as

short as 100 μm . There is always a possibility that there exists free-free absorption in a compact ionized region which has not yet been observed in emission, but still has a high covering fraction in most Seyfert 1 nuclei. Without observational evidence for such gas, it is unlikely that free-free absorption could be responsible for the observed steep sub-millimetre spectrum; but this possibility cannot be completely ruled out.

Thermal dust and non-thermal synchrotron emission remain to be considered. The steepest cutoff allowed by synchrotron self-absorption is $\alpha_{155-438\ \mu\text{m}} = +2.5$, so these observations can be used to place an upper limit on the flux at 155 μm arising from a synchrotron self-absorbed source, with the rest then arising from a thermal dust source. The interpretation is complicated, however, by the fact that the 155- μm image had a resolution of 45", whereas the 438- μm observations were made with a single 11" beam. Assuming that NGC4151 is at the mean distance of the Canes Venatici/Ursa Major group¹⁹ ($D = 18\text{ Mpc}$ for $H_0 = 75\text{ km s}^{-1}\text{ Mpc}^{-1}$), the 45" KAO and 11" JCMT apertures correspond to projected diameters of about 4 kpc and 1 kpc, respectively. To constrain the emission mechanism further, it is necessary to make an assumption about the spatial distribution and emissivity of the dust in NGC4151. A variety of different assumptions were explored, of which indicate that no more than about half of the 155- μm emission can be due to synchrotron self-absorbed emission. We present two possibilities here.

The first is that all the dust emission comes from within the inner 1 kpc. In this case, it is straightforward to interpret the limit on the sub-millimetre spectral index, $\alpha_{155-438\ \mu\text{m}} \geq +3.06$, as a limit to the slope of the dust emissivity law, $\epsilon \propto \lambda^\gamma$, corresponding to $\gamma \leq -1.06$. If $\epsilon \propto \lambda^{-1.5}$, corresponding to the standard mix of silicate and graphite grains⁶, less than 1.5 Jy (32%) of the flux density at 155 μm can be synchrotron self-absorbed emission. Even if $\epsilon \propto \lambda^{-2}$, which is the steepest emissivity law expected, less than 2.1 Jy (45%) of the 155- μm flux can be synchrotron self-absorbed emission.

A second approach is to assume that the dust is distributed uniformly across the 4-kpc KAO aperture. The KAO observations give some evidence of extended emission⁷, supporting this possibility. This yields an upper limit of $\leq 2.6\text{ Jy}$ ($\leq 54\%$) of the 155- μm emission which could emanate from a synchrotron self-absorbed source, essentially independently of the emissivity law.

These results indicate that cool dust emission dominates the sub-millimetre emission from NGC4151; however, different processes could be important in the near- and mid-infrared. A consistent explanation of all of the published infrared/sub-millimetre observations is that synchrotron emission dominates the near- to mid-infrared, cutting off at $\sim 80\ \mu\text{m}$, and that dust emission from star formation regions in the disk dominates in the sub-millimetre region, peaking at $\sim 110\ \mu\text{m}$. There are a number of arguments to support this picture: (1) After allowing for the 10.6–12- μm spectral slope and the difference in the calibration, the 60" IRAS 12- μm measurements and 6" ground-based 10.6- μm measurements agree to within 6%, so the mid-infrared source is quite compact⁴. However, the 155- μm source appears to be extended². (2) Quasar spectra tend to have higher frequency turnovers than Seyfert 1 galaxies, suggesting that the low-frequency turnover is masked by emission from cool dust in the galactic disks or narrow-line regions of lower-luminosity objects^{2,4} such as NGC4151. And (3), quasars ($\bar{\alpha}_{2.2-25\ \mu\text{m}} = -1.09$, $\bar{\alpha}_{12-60\ \mu\text{m}} = -0.69$) and luminous Seyfert 1 galaxies ($\bar{\alpha}_{2.2-25\ \mu\text{m}} = -1.15$, $\bar{\alpha}_{12-60\ \mu\text{m}} = -1.08$) tend to have relatively flat infrared spectra (presumably from non-thermal synchrotron emission), whereas Seyfert 2 galaxies ($\bar{\alpha}_{2.2-25\ \mu\text{m}} = -1.56$, $\bar{\alpha}_{12-60\ \mu\text{m}} = -1.55$) and other dusty objects have steeper spectra^{2,4}. NGC4151 has a relatively flat infrared spectrum⁴, with $\alpha_{2.2-25\ \mu\text{m}} = -1.31$ and $\alpha_{12-60\ \mu\text{m}} = -0.73$, so it appears that non-thermal emission dominates at near- and mid-infrared wavelengths.

The parameters derived for the thermally emitting dust are similar to those determined for other low-luminosity active galaxies and starburst galaxies. For a turnover of 110 μm and an emissivity law $\epsilon \propto \lambda^{-1}$, the dust temperature $T_d \approx 38\text{ K}$ (the temperature is only a weak function of the emissivity law). For a 110- μm flux density of $F \approx 8\text{ Jy}$, the derived dust mass $m_d \approx 10^5 M_\odot$, and the minimum size of the emitting region $r_d \geq 70\text{ pc}$. For a gas-to-dust ratio of 100, the mass of gas $m_g \approx 10^7 M_\odot$. These values are reasonable both for the narrow-line region and for star formation in the disk of the underlying galaxy.

It is important to try to extend the sub-millimetre observations of NGC4151 to higher frequencies and better sensitivities, in order to study the properties of the thermally emitting dust grains. These results show that the UKT14 system on the JCMT is sufficiently sensitive to detect, or obtain significant upper limits for, active galaxies. It is vital to expand the sample of objects observed at 438 μm , especially emphasizing luminous quasars with flat spectra and high turnover frequencies. These are the sources for which synchrotron self-absorbed emission is expected to dominate in the sub-millimetre region, and therefore only observations of such objects can provide the crucial test of the relative importance of different emission mechanisms. These objects are, unfortunately, also the most difficult to detect.

We thank J. M. Shull and M. Voit for helpful comments, and A. Lawrence and M. Elvis for allowing use of their data before publication. R.A.E. was supported by NASA, and M.A.M. was supported by NSF.

Received 5 July; accepted 11 November 1988.

1. Edelson, R. A. & Malkan, M. A. *Astrophys. J.* **308**, 59–77 (1986).
2. Edelson, R. A. *Astrophys. J.* **309**, L69–L72 (1986).
3. Carleton, N. et al. *Astrophys. J.* **318**, 595–611 (1987).
4. Edelson, R. A., Malkan, M. & Rieke, G. H. *Astrophys. J.* **321**, 233–250 (1987).
5. Rybicki, G. B. & Lightman, A. P. *Radiative Processes in Astrophysics* (Wiley-Interscience, New York, 1979).
6. Dwek, E. *Astrophys. J.* **322**, 812–822 (1987).
7. Engargiola, G. et al. *Astrophys. J.* **332**, L19–L22 (1988).
8. Orton, G. S. et al. *Icarus* **67**, 289–304 (1986).
9. Edelson, R. A. *Astr. J.* **94**, 1150–1155 (1987).
10. Clavel, J. et al. *Astrophys. J.* **321**, 251–279 (1987).
11. Cutri, R. et al. *Astrophys. J.* **296**, 423–429 (1985).
12. Rieke, G. H. & Lebofsky, M. J. *Astrophys. J.* **250**, 87–102 (1980).
13. Edelson, R. A. & Malkan, M. A. *Astrophys. J.* **323**, 516–535 (1987).
14. Edelson, R. A. *Astrophys. J.* **313**, 651–661 (1987).
15. Condon, J. J. et al. *Astrophys. J.* **246**, 624–646 (1981).
16. Pounds, K. A. et al. *Mon. Not. R. astr. Soc.* **218**, 685–694 (1985).
17. Kwan, J. & Krolik, J. *Astrophys. J.* **250**, 478–493 (1981).
18. Smith, M. G. et al. *Mon. Not. R. astr. Soc.* **195**, 437–449 (1981).
19. Geller, M. & Huchra, J. *Astrophys. J. Suppl.* **52**, 61–88 (1983).
20. Lawrence, A. et al. *Mon. Not. R. astr. Soc.* (submitted).

Quenching and disruption of lunar KREEP lava flows by impacts

Graham Ryder

Lunar and Planetary Institute, 3303 NASA Road 1, Houston, Texas 77058, USA

The Apollo 15 KREEP basalts—so called because they are rich in trace elements such as potassium, rare-earth elements and phosphorus—are thought to be volcanic rocks¹, based on their petrographic characteristics, chemical variations and geological context². The lavas were extruded $\sim 3.85\text{ Gyr ago}$ ³, and most are represented by only tiny fragments (nearly all less than 1 cm across). Here I report the results of a re-examination of their petrography. Several of them contain yellow residual glasses which cross-cut the crystallized phases; some show more extreme disruption. The features of the glasses appear to be compatible only with impact disruption, ejection and quenching from actively crystallizing flows, indicating a high impact flux immediately after the impact that formed the Imbrium basin. No other example of impacts into active lava flows is known in the Solar System.

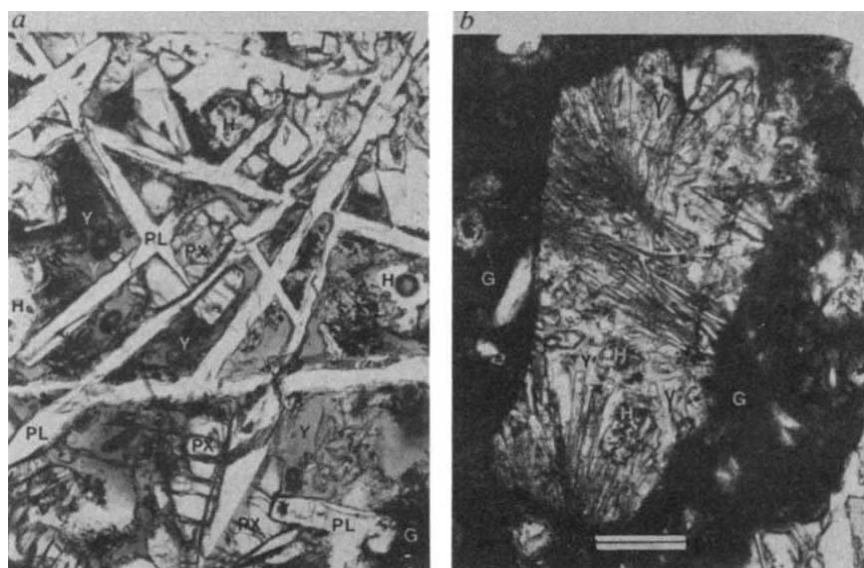


Fig. 1 Photomicrographs of two fragments in sample 15358 that contain yellow glass and are little disrupted. Plane-polarized light; scale bar represents $\sim 100 \mu\text{m}$ for both images. Yellow glass (Y, medium gray) is interstitial to plagioclase laths (PL, white, low relief) and pyroxenes (PX, white/pale gray, higher relief). H denotes holes in sections; G denotes impact-melt groundmass of 15358. The fragment shown in *b* contains very little pyroxene, which is not shown; most of its silicates are plagioclases (white), forming vague spherulites.

In the samples that contain yellow glass, most of the glass occurs interstitially, taking the place of the cryptocrystalline opaque interstitial material (mesostasis) that is characteristic of the glass-free fragments⁴. The textures are dominantly intersertal; two examples of contrasting grain size are shown in Fig. 1*a* and *b*. The glass-bearing fragments have a range of grain sizes and textures otherwise similar to other Apollo 15 KREEP basalts, except that the coarsest grain sizes are not represented. The yellow-glass-bearing samples have thermal histories that differ from those of the others only in that cooling is much faster in their late stage of solidification¹. The glasses are generally very clear, indicating quenching, which in turn suggests that the melt must have been situated close to a radiating or conducting surface. Sample 15358 consists of numerous varied KREEP basalt clasts in a fine-grained impact melt matrix (Fig. 1*b*); every clast in this sample has a yellow glass mesostasis. Several other samples occur as discrete coarse-fines (4–10 mm) and finer fragments (2–4 mm) in soils, and as clasts in other breccias. I have so far observed these yellow-glass-bearing basalts only among samples from the Apennine Front, and not from the mare plain. Although a few such samples have been observed before^{5,6}, their significance was not recognized.

In all fragments except one there has been disruption of the silicate minerals in the basalts, and movement of the yellow material into the fractures immediately before quenching (Fig. 2*a–c*). There is no visible or measured reaction of the veinlet glass with the silicates, but in a few instances it appears that ilmenite has been mildly resorbed by the moving melt. In some cases the yellow liquid became immiscible before moving and quenching, producing a small amount of Si, K-rich liquid. In all such cases, only the yellow glass moved and intruded the silicates (Fig. 2*b*); presumably the Si, K-rich liquid was too viscous to move under the same conditions. Two samples show extreme disruption (Fig. 2*c*), in which the silicate framework has been totally disrupted and the silicates now occur as clasts in a yellow-glass mesostasis; each of the two samples contains only mineral fragments derived from Apollo 15 KREEP basalts.

The yellow glasses, analysed by microprobe, have higher contents of TiO_2 , FeO, K_2O and P_2O_5 , and lower MgO, Al_2O_3 , Cr_2O_3 and SiO_2 , than the bulk KREEP basalts (Table 1). The chemical composition of the glasses varies amongst fragments (Table 1, Fig. 3), but is fairly constant within a sample, with the exception of two whose silicates are fine-grained (open circles in Fig. 3), and for these it appears that local pockets of liquid were cut off from each other and crystallized differently. The chemical compositions of the yellow glasses are consistent

with their being residual liquids resulting from the crystallization of Ca-plagioclase and low-Ca pyroxenes, and later ilmenite, from Apollo 15 KREEP basalts. Within a sample there is no difference in composition between veins of glass which cut silicates and glass which remains interstitial. The compositional differences of the glasses amongst fragments correlates with other differences: the less glass there is in a sample, the more evolved it is (higher TiO_2 , K_2O , FeO and so on), the more Fe-rich is the pyroxene, and the more Na-rich is the plagioclase that last crystallized before quenching occurred. The least-evolved glasses did not crystallize either cristobalite or ilmenite, nor had they become immiscible; more-evolved glasses crystallized successively cristobalite, then cristobalite and ilmenite; the most-evolved glasses had crystallized these phases and had also developed immiscibility. These features are summarized in Fig. 3, and they all demonstrate that the glass is intimately related to the mineralogical features of the host fragment and is not

Table 1 Major-element compositions of a typical Apollo 15 KREEP basalt and selected residual yellow glasses

	Sample						
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>
SiO_2	51.4	50.2	48.8	47.3	46.4	46.8	45.2
TiO_2	2.2	3.4	5.7	5.0	5.3	5.4	4.7
Al_2O_3	15.7	12.1	10.6	9.0	9.2	9.4	7.9
Cr_2O_3	0.31	0.28	0.07	0.19	0.11	0.07	0.08
FeO	10.0	13.2	15.4	17.5	18.1	20.2	22.3
MnO	0.15	0.19	0.19	0.23	0.25	0.28	0.29
MgO	8.0	6.7	2.8	4.2	3.7	2.9	2.8
CaO	9.8	9.6	9.6	10.1	10.9	10.5	10.0
Na_2O	0.79	0.53	0.18	0.50	0.43	0.42	0.49
K_2O	0.60	0.91	1.3	1.2	1.2	1.3	1.4
P_2O_5	0.56	1.2	2.0	2.1	2.7	2.1	3.3

a, 15434,16 whole-rock composition, Apollo 15 KREEP basalt. *b–g*, yellow glasses: *b* 15243,16, totally disrupted sample; *c*, 15358,6 clast, no disruption; *d*, 15434,29, totally disrupted sample; *e*, 15434,21, minor disruption; *f*, 15358,16 clast, moderate disruption, immiscibility; *g*, 15434,25 strong disruption, immiscibility. All data are from this work. The KREEP basalt is from microprobe analysis of a fused glass bead of a powder made by grinding a whole-rock sample, and the yellow-glass data are averages of multiple microprobe analyses of glasses in thin sections. All analyses were made using the Cameca microprobe at the Johnson Space Center. For all elements, the 1σ precision of single point counts is $<2\%$ (commonly $<1\%$) except for elements present at $<1 \text{ wt}\%$. For averages, the standard deviations for single elements are 3–4% for most elements, rising to 5–7% for K, Na, Mg and Si.

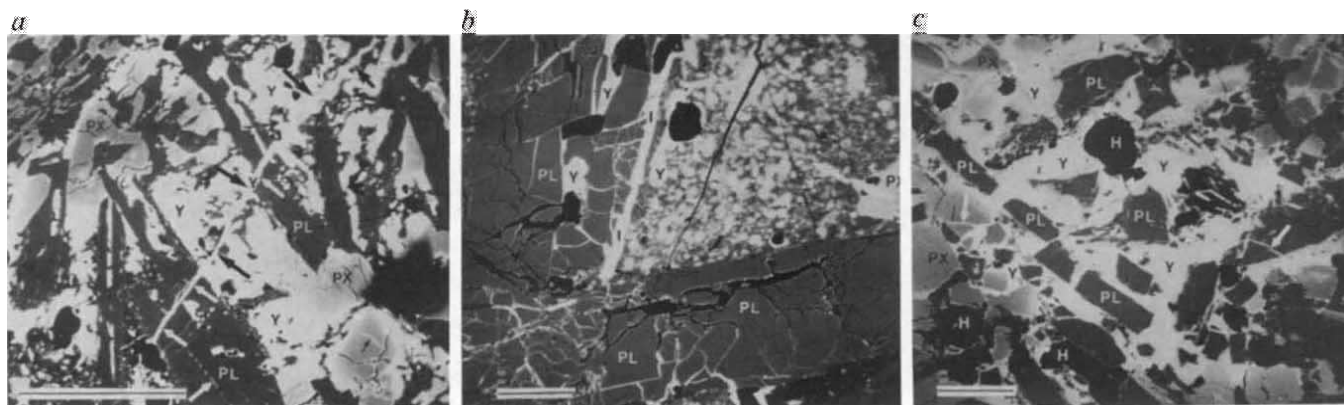


Fig. 2 Backscattered-electron images (emphasizing composition) of disrupted fragments of KREEP basalts containing yellow glass. Scale bars on each image represent 200 μm . Y (white/very pale gray) denotes yellow glass. PL (dark gray) denotes plagioclase. PX (pale gray) denotes pyroxene. H (black) denotes vesicles or plucked holes. *a*, Sample 15434,21. Little disrupted, with minor veins cutting the silicates. A prominent yellow-glass vein is indicated by the arrows and runs from near top right to near bottom left. Other arrows indicate miniscule veins. Part of the mesostasis has crystallized or devitrified (mottled gray patches). *b*, Sample 15434,25. More disrupted. Black areas are holes and unfilled fractures. A prominent triangular mesostasis area (top right) has undergone immiscibility to produce globules of yellow glass (white) and silicic areas (gray). Only the yellow material intruded the two prominent fractured plagioclase crystals; the veining of these plagioclases is pervasive. *c*, Sample 15434,29. Totally disrupted. Fragments of minerals have not all moved very far, as shown by broken plagioclase in left centre, indicating that quenching followed disruption very rapidly. Most of the yellow glass (very pale gray) forms a massive groundmass; the arrows point to yellow glass that veins fractured silicates.

introduced; that is, the glasses have the chemical features expected of a quenched residual liquid appropriate for the sample in which they are found. This is true for the two totally disrupted samples as well as for the others. The liquid line of descent is towards silica depletion and iron-enrichment, not towards 'granitic' compositions.

I have determined the whole-rock compositions of several of the yellow-glass-bearing fragments by instrumental neutron activation analysis (INAA), and in some cases by microprobe fused-bead techniques⁷ at the Johnson Space Center. For others the compositions have been estimated from the glass and mineral proportions and compositions. All of the yellow-glass-bearing samples have major- and trace-element whole-rock compositions in the range of normal Apollo 15 KREEP basalts with cryptocrystalline mesostases. No meteoritic contamination occurs in any of them above the detection limits (~ 3 parts per 10^9 for Au). This is true also for the two totally disrupted samples. In no case is the whole-rock composition similar to the yellow glasses, which are best interpreted as residual compositions.

The yellow glasses cannot be produced by total impact-melting of KREEP basalts (which would explain the movement), because they have the compositions expected of residual melts of the host fragment in which they are found, and not of whole-rock compositions, as would be expected of impact melts⁸. They are also unlikely to be caused by auto-brecciation, in which solidified crust is stressed and then broken by continued movement of lava⁹. This mechanism fails here because (1) it is a slow process that does not intrude veinlets a few micrometres wide over distances of several millimetres; such veinlets occur in some of the yellow-glass-bearing samples that otherwise show no evidence of deformation, and require a more rapid and forceful mechanism; (2) it moves material that is more like the original basalt than a local residual liquid; (3) it does not explain the totally disrupted samples, and because it is a continuous process one would expect it to result in some samples showing multiple episodes, whereas as far as I can tell, none do; (4) it cannot explain those rare samples with quenched glass that lack movement of the glass or disruption of the crystals; (5) if auto-brecciation occurred, one would also expect some glass veins to cut normal samples with cryptocrystalline mesostases, but none have been observed to do so; (6) sample 15358, containing exclusively fragments with yellow glass, would be

an unlikely consequence of auto-brecciation; one would expect some non-glassy fragments to have been in the target for the impact that is then required to produce the 15358 groundmass.

A mechanism (suggested first by Phinney *et al.*⁵ for all the mesostases, whether yellow glass or cryptocrystalline) that involves partial melting of interstitial material has been proposed by O. B. James (personal communication), using impact heating of some sort to cause the melting. This mechanism fails here because (1) it provides no way of moving the melt after it has formed; the movement required impact after melt formation; (2) it provides no way to quench the melt to glass; (3) it cannot explain the two totally disrupted samples; (4) increasing temperature does not produce silicate liquid immiscibility; (5) it would not preserve the euhedral crystals, such as ilmenite needles, observed in the yellow glass of many samples; (6) it would predict regular gradients of melting across samples, including preservation of some opaque mesostasis areas, whereas none such occur; such gradients do occur, however, in known examples of melting produced adjacent to splashed glass coats in lunar samples of a scale similar to those I have inspected; and (7) it cannot explain the abundance of the varied fragments found together in sample 15358, all containing yellow glass to the absolute exclusion of other basalt fragments with more typical textures.

All of these known features may, however, be explained by a mechanism involving impacts into a crystallizing flow(s) or lake(s) of Apollo 15 KREEP basalts, splashing out liquids and crystal/liquid mushes, some of which were ejected to the Apennine Front and preserved. Such impacts would disrupt the flow and quench the residue at varied stages of evolution. This mechanism explains the quenching of the yellow glass, the fracturing of the silicates and the movement of the glass immediately before quenching, the existence of the two totally disrupted samples, and the varied stages of liquid evolution, including silicate liquid immiscibility. This process also can produce samples that show no evidence of disruption but are merely quenched. Furthermore, it can explain the existence of sample 15358, as a total impact-melt with meteoritic contamination (the fine melt groundmass of 15358) and clasts of basalt from immediately around the target which were caught up after quenching. Most of the residual yellow glasses in the 15358 clasts are rather similar in composition, despite the variety of textures. The

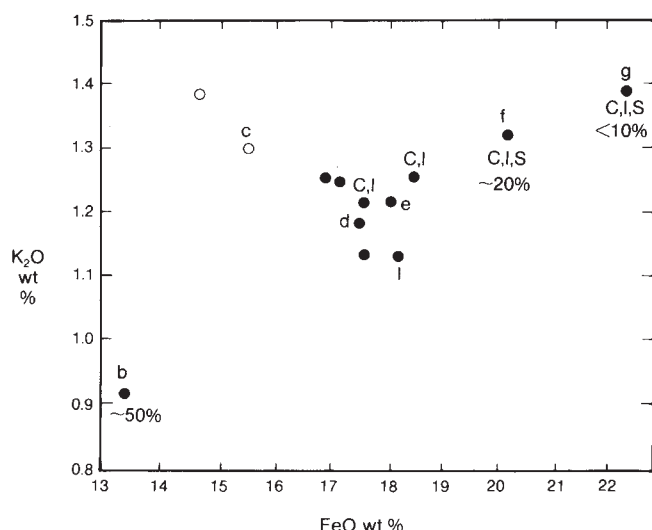


Fig. 3 K_2O content plotted against FeO for average yellow glass in samples, with superimposed information for each sample indicating % glass, presence of cristobalite (C), ilmenite (I), and presence of silicate liquid immiscibility (S). The data form a sequence from the least-evolved residue (earliest quench) at the left to the most-evolved residue (latest quench) at the right. Lower-case letters correspond with analyses in the top column of Table 1. Open symbols are for two fine-grained samples with much more varied glass analyses, which cannot be considered as necessarily good averages.

varieties of glass compositions amongst other fragments suggests several impacts at different times or a single impact into a flow which had locally varied stages of evolution (which is possible in, for example, a crystallizing lava lake). A prediction of this hypothesis is that all of these samples, including the yellow glasses themselves and the groundmass of 15358, have the same ages as the other Apollo 15 KREEP basalts analysed, specifically 3.85 Gyr. This remains to be tested. The process suggested would not have incorporated any siderophiles into the samples containing yellow glass; the role of any impact was to shock and eject material that already existed. Nor would the flux necessarily be high enough that the remaining KREEP flows themselves would be significantly affected by siderophile addition; the impactors would still be a very tiny fraction of the total material in the flows.

Even if only one impact produced the entire range of samples observed, which were from different sampling locations, one must infer either a fortuitous occurrence or an impact flux that was very high during the extrusion of these basalts. There are no similar examples among mare basalts, which have been collected in much greater numbers. Indeed, no other example of meteorite impact into a crystallizing flow is known from the Solar System. It is possible that the generation of Apollo 15 KREEP basalts was induced by the impact that created the Imbrium basin; their formation certainly immediately followed it^{10,11}. The evidence presented here suggests that the flux of small objects remained high immediately after the Imbrium event, which was consequently not a single large isolated event. This provides a dynamic constraint on the origin of the 3.85-Gyr surface events.

This work was conducted under NASA grants. The Lunar and Planetary Institute is operated by the Universities Space Research Association under contract with the National Aeronautics and Space Administration. I appreciate the help of personnel in the Johnson Space Center neutron activation and microprobe laboratories.

Received 1 September; accepted 7 November 1988.

1. Ryder, G. *Proc. Lunar planet. Sci. Conf. 17th*, in *J. Geophys. Res.* **92**, E331-E339 (1987).
2. Spudis, P. D. & Ryder, G. *Eos* **66**, 724-726 (1985).
3. Carlson, R. & Lugmair, G. *Earth planet. Sci. Lett.* **45**, 123-132 (1979).
4. Dymek, R. *Lunar Planet. Inst. Tech. Rep. 86-03* (eds Spudis, P. D. & Ryder, G.) 52-57 (1986).
5. Phinney, W., Warner, J., Simonds, C. & Lofgren, G. *The Apollo 15 Lunar samples* (eds Chamberlain, J. W. & Watkins, C.) 149-153 (Lunar Science Institute, Houston, 1972).
6. Crawford, M. & Hollister, L. *Proc. Lunar Sci. Conf. 8th* 2403-2417 (1977).
7. Ryder, G. *51st A. Mtg. met. Soc. Fayetteville, Arkansas*, H-9 (1988).
8. Phinney, W. & Simonds, C. *Impact and Explosion Cratering* (eds Roddy, D. J., Pepin, R. O. & Merrill, R. B.) 771-790 (Pergamon, New York, 1977).
9. Macdonald, G. A. *Volcanoes* 510 (Prentice Hall, New Jersey, 1972).
10. Papanastassiou, D. & Wasserberg, G. *Proc. Lunar Sci. Conf. 7th* 2035-2054 (1976).
11. Ryder, G. *51st A. Mtg. met. Soc. Fayetteville, Arkansas*, H-4 (1988).

The formation of rain in convective clouds

Anthony J. Illingworth*

NCAR, PO Box 3000, Colorado 80307, USA

One of the chief problems in cloud physics¹ is to explain how, in the absence of ice, small cloud droplets can coalesce to form raindrops. Cloud droplets are produced in concentrations of $\sim 10^8 \text{ m}^{-3}$ by condensation on sub-micrometre atmospheric nuclei, but, because they are so small, collisions between droplets are rare events. In mature rainclouds there are² usually about 1,000 drops per m^3 ; here, however, I present differential radar reflectivity observations which indicate that, when ice is not present, first echoes of convective showers consist of large ($>4 \text{ mm}$) raindrops present in much lower concentrations ($<1 \text{ m}^{-3}$). I suggest that these raindrops form on ultra-giant nuclei ($>30 \mu\text{m}$ radius) which are present in similarly low background concentrations³. In some clouds this initial raindrop spectrum persists, whereas in other cases it changes rapidly and irreversibly to the raindrop distribution normally observed in mature clouds with many more smaller raindrops. The latter may occur when the break-up of the large raindrops becomes more common and many small satellite drops are produced.

Coalescence of small droplets in clouds may be triggered by condensation on hygroscopic salt nuclei or by mixing of drier air into the cloud; either process could produce larger droplets, which would then have a reasonable terminal velocity and be able to capture smaller droplets, thus triggering droplet growth. Paluch and Knight⁴ have reviewed the evidence for these processes. The cloud droplet concentration is at least five orders of magnitude higher than the raindrop concentration, so that it is very difficult to make direct in-cloud observations of the occasional embryonic raindrop in the radius range 50-100 μm , when so many smaller droplets are also present. Measurements of the conventional radar reflectivity, Z , are difficult to interpret in terms of both the size (d) and concentration (n) of raindrops, because Z is given by the product nd^6 summed over all raindrop sizes. Z is usually quoted in dBZ, that is in decibels over $1 \text{ mm}^6 \text{ mm}^{-3}$. A full review of the meteorological applications of radar is provided by Browning⁵.

Here we report measurements of differential reflectivity (Z_{DR}), which is defined as

$$Z_{DR} = 10 \log (Z_H/Z_V) \quad (1)$$

where Z_H and Z_V are the radar reflectivity factors measured with horizontally and vertically polarized radiation, respectively. Z_{DR} is a measure of the shape of precipitation particles. Because raindrops are oblate to a degree which depends on their size, Z_{DR} is positive for large raindrops and its magnitude is a measure of their mean size. With knowledge of the drop size, d , the drop concentration n may be calculated⁶ from the absolute value of

* Permanent address: Department of Physics, UMIST, Manchester M60 1QD, UK.

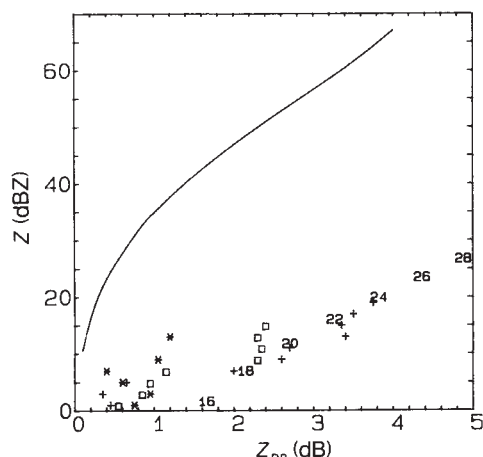


Fig. 1 Values of Z (radar reflectivity) and Z_{DR} (differential reflectivity) observed during the evolution of the echo on 10 July 1986. The solid line represents the drop concentrations for average rain, $n_0 = 8,000 \text{ m}^{-3} \text{ mm}^{-1}$ (equation 2). The observations are average values of Z_{DR} for each 2-dBZ step in Z . Symbols represent observation times: * 12:57, □ 13:01, + 13:04. n_0 scales linearly with Z , so inferred raindrop concentrations are over 30 dB (a factor of one thousand) below 'average'. The numbers indicate time in minutes elapsed since the initial evolution of the radar responses Z and Z_{DR} , during which time ultra-giant nuclei sweep out cloud water droplets.

Z_H . In practice the raindrop size distribution may be well approximated by a spectrum of the form:

$$n(d) = n_0 \exp(-3.67d/d_0) \quad (2)$$

where d_0 is the equivolumetric diameter. The value of Z_{DR} is independent of concentration and may be calculated as a function of d_0 . The value of n_0 may then be found from the observed magnitude of Z . The solid curve in Fig. 1 shows calculated values of Z and Z_{DR} for $n_0 = 8,000 \text{ m}^{-3} \text{ mm}^{-1}$; this value of n_0 was originally proposed for stratiform rain² but it now appears to describe the average properties of most rain. For individual data points and clouds not lying on this average curve, the value of n_0 may be derived by noting that Z is proportional to n_0 . The total concentration is given by $n_0 d_0 / 3.67$.

Observations of developing convection were made each day during the MIST project⁷ in 1986 using the 10-cm dual polarization CP-2 radar operated by the National Center for Atmospheric Research. The values of Z and Z_{DR} reported here are accurate to better than 1 dBZ and 0.2 dB respectively. The first echoes of 5–10 dBZ usually formed at temperatures warmer than 0 °C and were accompanied by anomalously positive values of Z_{DR} , implying (Fig. 1) the presence of a very low concentration of unexpectedly large raindrops and confirming more limited UK observations^{8,9} made in 1983 and 1984 with the Chilbolton polarization radar.

A particularly well observed case occurred on 10 July 1986, when a series of 90 vertical (RHI) and horizontal (PPI) radar sections was made through an isolated echo from its initiation with $Z = 5$ dBZ at 12:54 CST (=GMT–6 hours) to a maximum of $Z = 20$ dBZ (accompanied by values of Z_{DR} near 4 dB), and terminating at 13:10 as the echo fell to the ground and disappeared. During this time the echo was unsheared, its top never rose above 5 km altitude and the horizontal cross-section never exceeded 2 km in diameter. It is not possible to plot the hundred or so data points obtained for each scan, but a summary of the evolution is provided in Fig. 1, in which the average value of Z_{DR} for each 2-dBZ step in Z is plotted for three scans at 12:57, 13:01 and 13:04. These scans are all at the same azimuth and pass through the maximum echo. We note that the maximum

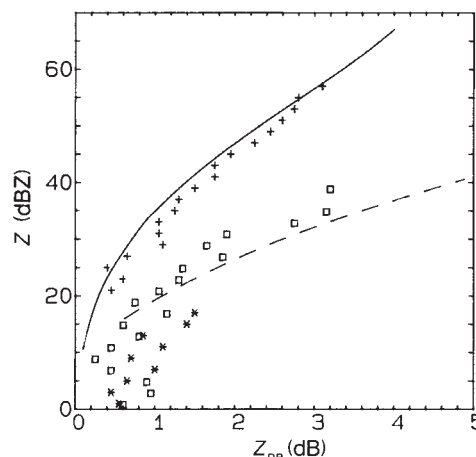


Fig. 2 The evolution of the echo observed on 28 June 1986. Solid line and averages follow the same key as in Fig. 1. Times are: * 12:21, □ 12:27, + 12:40. The dashed line joins the values of Z and Z_{DR} that correspond to a 5-mm raindrop which has a lifetime of 1,000 s before shattering as a result of collisions. For a given Z_{DR} the lifetime is inversely proportional to Z .

values of Z and Z_{DR} increase with the passage of time, but that the average values all lie on a similar curve. Plots for the intermediate scans show a smooth progression along this evolutionary curve. Increases in Z ceased at 13:04, as the main echo fell below cloud base (inferred from a nearby sonde ascent) and the plots for the scan at 13:07 are within 0.5 dB of the 13:04 data.

The evolution shown in Fig. 1 is compatible with a situation in which a very few raindrops grow by sweeping out cloud liquid water. In 3 min the maximum Z_{DR} increased from 2.4 dB to 4 dB, corresponding to an increase of the maximum diameter for monodispersed drops from ~4 mm to 6 mm, this occurring by collection of cloud droplets with a liquid water content of 2 g m^{-3} . The concentrations are very low; for example, a Z value of 15 dBZ accompanied by a Z_{DR} of 3 dB is equivalent, for monodispersed drops, to the presence of 4.5-mm drops at a concentration of 0.03 m^{-3} , or, for an exponential distribution, to a value of n_0 of $0.5 \text{ m}^{-3} \text{ mm}^{-1}$. The difference in concentrations predicted by the exponential and the monodispersed spectra is negligible⁹ compared with the deviation from the 'average' curve in Fig. 1. Analysis of the sonde ascent predicts a maximum adiabatic liquid water content of 3 g m^{-3} for a temperature of 8 °C. Above this level, much drier air, with a mixing ratio of less than 2 g m^{-3} at 5 °C, restricted cloud growth.

The data are in very good agreement with the time evolution predicted by a model⁸ in which ultra-giant nuclei collect cloud water and grow to raindrops. This model extends earlier work¹⁰ which showed that such a mechanism could produce a high radar reflectivity within 15–20 min. Our new model predicts the accompanying values of Z_{DR} , but does not consider the vertical structure that may result from sedimentation. The nuclei need not be hygroscopic but, because they are so large, they collect the cloud droplets with reasonable efficiency. The numbers in Fig. 1 indicate the time in minutes elapsed since initiation of the mechanism proposed in our model, for a cloud liquid water density of 2 g m^{-3} with concentrations (in units of m^{-3} per unit log interval of radius) of 1,000 for nuclei of 10- μm radius to 0.03 for 100- μm radius. These nuclei concentrations are those quoted by Junge³ for the background level found above the trade-wind inversion and are similar to the 'maritime' nucleus spectrum used by Johnson¹⁰. The drop concentrations inferred on 10 July are exceptionally low; values a factor of ten higher were inferred on other days in the MIST programme. It may be that on such days nucleus concentrations closer to the ground

were higher than the background levels used in the model.

When the convection was vigorous, the intensifying echo top usually rose above the 0 °C isotherm and the spectra in the rain tended towards that usually observed in mature clouds, but on one occasion (28 June 1986) it was possible to observe the evolution of an echo to over 50 dBZ without ice forming in the cloud. The average values of Z_{DR} for each 2-dBZ step in Z for three successive scans are shown in Fig. 2. The time resolution is rather poor because the radar was executing 360° surveillance scans. At 12:21 the maximum Z value was 17 dBZ, associated with a Z_{DR} value of ~1.5 dB, and inferred concentrations are more than two orders of magnitude (equivalent to 20 dB) below the curve for 'average' rain. Six minutes later, at 12:27, Z had reached nearly 40 dBZ and Z_{DR} was ~3 dB, indicating that the biggest drops had grown in diameter by 2 mm. The concentration of these large drops is still two orders of magnitude below average. By 12:40 the highest Z was over 55 dBZ, but the picture is then dramatically different, the inferred raindrop size distribution being that found in mature clouds.

To support our contention that drop collisions are responsible for this abrupt change in the spectra, the dashed line in Fig. 2 represents values of Z and Z_{DR} for which a 5-mm droplet would have an average lifetime of 1,000 s before it collided with another droplet greater than 1 mm in diameter. Collisions between two raindrops which are both larger than 1 mm are generally followed by disruption and shattering to produce many small fragments¹¹. For a given Z_{DR} , the value of d_0 is constant and the number of drops is proportional to Z ; accordingly, the lifetime of each large raindrop is inversely proportional to Z . We note that in Fig. 1 the data lie in a region where the large drops have a lifetime of several hours before they suffer collision-induced rupture. In Fig. 2 the initial concentrations are somewhat higher, and by 12:27 the drop lifetimes are slightly shorter than 20 min; finally, at 12:40 h the spectrum has changed to the 'mature' mode, with a high concentration of smaller drops, so that the large drops contributing to a Z_{DR} of 3–4 dB would have a lifetime of only a few seconds and would be continuously regenerated.

Some measurements^{12,13} made from aircraft within developing clouds have also indicated raindrop spectra different from those found in mature clouds, although Z values were higher and the spectra were not as extreme as those reported here. Measurements in Texas¹² revealed low values of n_0 in early echoes and an implied absence of collision-induced break-up. The authors attributed the low values of n_0 to drop sorting in the updraft, but this does not seem compatible with the time evolution of the Z_{DR} RHI scans from the MIST data. In Hawaii¹³ raindrops 8 mm in size were observed. The existence of such large raindrops had been questioned because in Marshall–Palmer rain they would rapidly suffer collision-induced break-up, but the authors showed that when total drop concentrations are low, the lifetimes are sufficiently long. The very low concentrations inferred from the radar data of the MIST project would be difficult to confirm by aircraft penetrations because of the limited sample volume of most precipitation-sizing probes.

Other mechanisms for the production of the raindrops must be considered. There is always the possibility that a few ice crystals (which would have a Z value below –20 dBZ and would be undetectable by radar) could fall from higher levels, melt, and then grow into large raindrops. This seems unlikely: the air was very dry at these colder levels and no bright band in Z_{DR} was observed. (It is our experience that as small ice crystals start to melt there is a sudden jump in Z_{DR} as they become wet and change their dielectric constant, followed by an equally sudden drop to zero Z_{DR} as they melt completely to become small spherical raindrops.)

In his initial report on the existence of ultra-giant nuclei, Junge³ remarked that they were so rare that they could not be important in the warm-rain process. Such a view seems eminently sensible, but this new radar evidence shows that warm rain does

indeed first form in such low concentrations. It may be that a fortunate combination of collection efficiencies and droplet sizes results in just one cloud droplet in 10^8 being able to capture its neighbours and grow into a raindrop. It is evident that further modelling and observations are needed, but I propose meanwhile that cloud droplet capture by ultra-giant nuclei is a simple and equally plausible mechanism.

I thank Professors Bringi, Forbes, Fujita and Wakimoto for allowing access to the MIST data, and staff at NCAR for their assistance during my visit. This research was also supported by NERC and EOARD. NCAR is supported by the National Science Foundation.

Received 2 August; accepted 14 November 1988.

1. Pruppacher, H. R. & Klett, J. D. *Microphysics of Clouds and Precipitation* (Reidel, Dordrecht, 1978).
2. Marshall, J. S. & Palmer, W. McK. *J. Met.* **5**, 165–166 (1948).
3. Junge, C. E. *J. geophys. Res.* **77**, 5183–5200 (1972).
4. Paluch, I. R. & Knight, C. A. *J. Atmos. Sci.* **43**, 1994–1998 (1986).
5. Browning, K. A. *Rep. Prog. Phys.* **41**, 761–806 (1978).
6. Seliga, T. A. & Bringi, V. N. *J. appl. Met.* **15**, 69–76 (1976).
7. Dodge, J., Arnold, J., Wilson, G., Evans, J. & Fujita, T. T. *Bull. Am. met. Soc.* **67**, 417–419 (1986).
8. Caylor, I. J. & Illingworth, A. J. *Q. Jl R. met. Soc.* **113**, 1171–1191 (1987).
9. Illingworth, A. J., Goddard, J. W. F. & Cherry, S. M. *Q. Jl R. met. Soc.* **113**, 469–499 (1987).
10. Johnson, D. B. *J. Atmos. Sci.* **39**, 448–460 (1982).
11. Lowe, T. B. & List, R. *J. Atmos. Sci.* **39**, 1591–1618 (1982).
12. Carbone, R. E. & Nelson, L. D. *J. Atmos. Sci.* **35**, 2302–2314 (1978).
13. Beard, K. V., Johnson, D. B. & Baumgardner, D. *Geophys. Res. Lett.* **13**, 991–994 (1986).

Deforestation alters denitrification in a lowland tropical rain forest

G. Philip Robertson* & James M. Tiedje†

* W. K. Kellogg Biological Station and Department of Crop and Soil Sciences, Michigan State University, Hickory Corners, Michigan 49060–9516, USA

† Department of Microbiology and Department of Crop and Soil Sciences, Michigan State University, East Lansing, Michigan 48824, USA

Nitrogen gas loss from terrestrial ecosystems is the most poorly quantified component of the global nitrogen cycle^{1–3}. Particularly little is known about gas losses from tropical rain forests, which may be an especially important source of nitrogen gases worldwide^{3,4}, and one that is changing rapidly because of tropical deforestation. Here we report measurements of denitrification ($N_2 + N_2O$ production), an important mechanism of nitrogen loss in most ecosystems^{5–7}, at a set of sites in Central America. Measurements were made to determine whether nitrogen gas loss is related to the successional status of rain-forest vegetation. Denitrification is high in primary forest and early successional sites and substantially lower in mid-successional sites. This implies that denitrification can be a major route of nitrogen loss from recently cleared and primary forest sites, and that global denitrification losses from humid tropical regions today are probably much smaller than losses in pre-colonial times, when a smaller proportion of sites were in mid-successional growth phases.

Recent work in temperate forests suggests that rates of denitrification may be closely related to the successional status of the existing plant community^{8,9}. Results suggest that nitrogen gas losses from old-growth and early successional communities may be much greater than from communities in which much of the vegetation is in a nitrogen-immobilizing, mid-successional growth phase. This pattern, if true, could have important consequences for the management of soil nitrogen resources in the humid tropics. It could also have important consequences for estimates of the impacts of human activity on global nitrogen gas fluxes¹⁰: if denitrification is an important source of N_2O in these ecosystems^{5–7}, then the worldwide loss of primary tropical

rain forests may be attenuating rather than contributing to some of the recently documented increases in atmospheric concentrations of nitrogen trace gases^{11,12}.

Over the past two years we have measured denitrification rates in rain-forest soils at several sites at the La Selva field station of the Organization for Tropical Studies in the wet Atlantic lowlands of north-east Costa Rica, Central America (10°26' N, 84°00' W, 50 m elevation). Rainfall at La Selva is weakly seasonal (mean annual precipitation is 4,000 mm) and temperature is relatively constant year-round, with an annual mean of 24 °C. Soils at the station vary from relatively fertile entisols developed on recent alluvial deposits of volcanic origin to highly weathered, base-poor inceptisols and ultisols derived from alluvium and lava flows¹³⁻¹⁵; all of our sample sites were on the older soils, dominated by low-charge-density, variable-charge clays. Three of the sites were on typic (site SWT) and andic (sites WP, SOC) Humitropept soils; the remainder were on oxic (sites GC, GBF, FBF) and andic (sites BP, PAS) Dystropepts¹³. Soils from all sites exhibited low pH (4.2–5.0), were well aggregated, and nitrified readily (2–4 $\mu\text{g g}^{-1} \text{d}^{-1}$ $\text{NO}_3^- \text{N}$) in 10-day potential mineralization assays (G.P.R. and J.M.T., unpublished results).

We measured denitrification at ten sites: four primary-forest sites which had never been cleared, two mid-successional sites cleared from primary forest about 25 years before sampling, two mid-successional sites cleared from successional vegetation about two years before sampling, and two early-successional sites cleared from successional vegetation several weeks before sampling. The maximum distance between sites was 2 km. Our primary-forest sites were typical for La Selva, where in intact forest a very diverse mixture of species form a 30–40-m-high canopy with 50-m emergents¹⁶. The two older mid-successional sites were on sites that had been first cleared of primary forest at around 1970; one site (PAS) was, at the time of sampling, a 2-m-high stand of grass and woody shrubs; the other site (GC) was a mixture of 2-m-high grass and fern. The two younger mid-successional sites (GBF and FBF) also had been cleared of primary forest around 1970, but in 1984, 18 months before their mid-successional sampling dates, they had also been cleared of regrowth (2-m-high grass and grass plus fern, respectively) and were maintained vegetation-free thereafter. Several weeks after their 1984 clearing, these two sites were our early-successional sites (early GBF and early FBF, respectively). All four of the mid-successional sites had similar pre-1984 land-use histories in that all were cleared around 1970, grazed for 5–8 years, and then abandoned. Rates of re-vegetation on these sites have been slower than those reported for other wet tropical sites in Costa Rica¹⁷; this is probably because our sites are on highly weathered, base-poor soils and, after cutting, were grazed for a number of years before abandonment.

Denitrification rates were measured by an aerobic, short-term intact-core assay¹⁸. On any given sample date (see below), 10–20 soil cores of 2.2-cm diameter $\times$ 15-cm depth were removed from each site at random intervals along three 10-m transects within a 1-hectare area. Cores were collected in individual acrylic tubes, which were then taken to a field laboratory, flushed with a (10 kPa $\text{C}_2\text{H}_2 + 90$ kPa air) atmosphere, and stoppered. N_2O accumulation over the ensuing 12 hours was monitored in each tube at three-hour intervals by injecting headspace gas into a gas chromatograph fitted with a ^{63}Ni electron-capture detector. Denitrification activity ($\text{N}_2 + \text{N}_2\text{O}$ production) was calculated from the linear portion of the N_2O accumulation curve^{19,20}. Rates measured in this manner have been shown to directly reflect *in situ* denitrification rates in temperate-region forest, grassland and agricultural sites^{21,22}; we assume that the same is true for our tropical site.

Sample dates varied somewhat between sites, with all except the early-successional sites sampled 2–3 times at 4–6-month intervals over the course of a year. The early-successional sites were sampled twice in the 6-month interval following clearing.

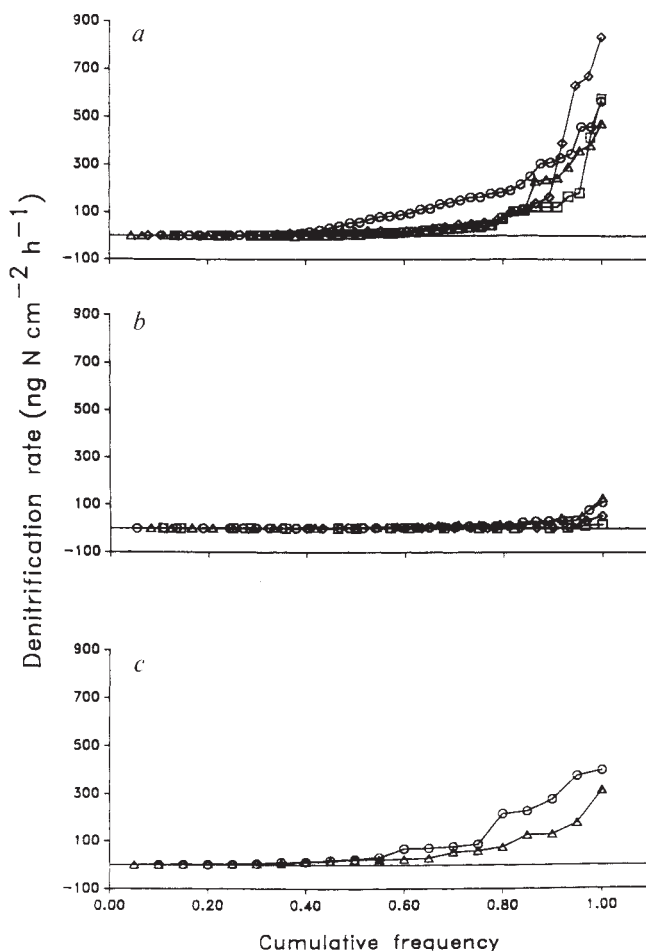


Fig. 1 Cumulative frequency distributions for denitrification rates ($\text{N}_2 + \text{N}_2\text{O}$ production) in individual soil cores from the 8 rain-forest sites sampled. Primary-forest sites have never been cut; all four mid-successional sites had been cut about 15 years before sampling, and two sites (GBF and FBF) had been cut again 1.5 years before sampling and maintained vegetation-free (vf). Values for the early-successional sites FBF and GBF represent samplings at 1 and 6 months following clearing. *a*, Primary-forest sites: $\circ$ WP, $\triangle$ SOC, $\square$ SWT, $\diamond$ BP. *b*, Mid-successional sites: $\circ$ PAS, $\triangle$ GC, $\square$ GBF (vf), $\diamond$ FBF (vf). *c*, Early-successional sites: $\circ$ early GBF, $\triangle$ early FBF.

Preliminary experiments suggested that individual rainfall events had no discernible effect on denitrification activity in these moist, extremely well-drained soils; nevertheless, no sites were sampled within 12 hours of the most recent rainfall. On the last sample date, the effects of added nitrate on denitrification were evaluated by adding either 0 or 100 $\mu\text{g NaNO}_3\text{-N}$ in 10 ml H_2O to each core and monitoring N_2O accumulation as above.

Within any given site, denitrification rates were quite variable even within individual sample dates. Rates among cores from within a site tended to vary by an order of magnitude or more, with frequency distributions of rates tending towards the log-normal (Fig. 1). This was true regardless of whether maximum rates within the site were low (for example, the mid-successional PAS site, with a mean rate of 3.61 $\text{ng N cm}^{-2} \text{h}^{-1}$) or high (for example, the primary-forest WP site, with a mean rate of 24.3 $\text{ng N cm}^{-2} \text{h}^{-1}$), and is typical of rates in many temperate-region sites^{9,23}. Partly because of this high within-site variability, there were no significant differences between different sample dates for any given site, and the date-by-site interaction term was negligible (probability $P > 0.10$). Consequently, for each site we pooled data across dates for subsequent statistical analyses.

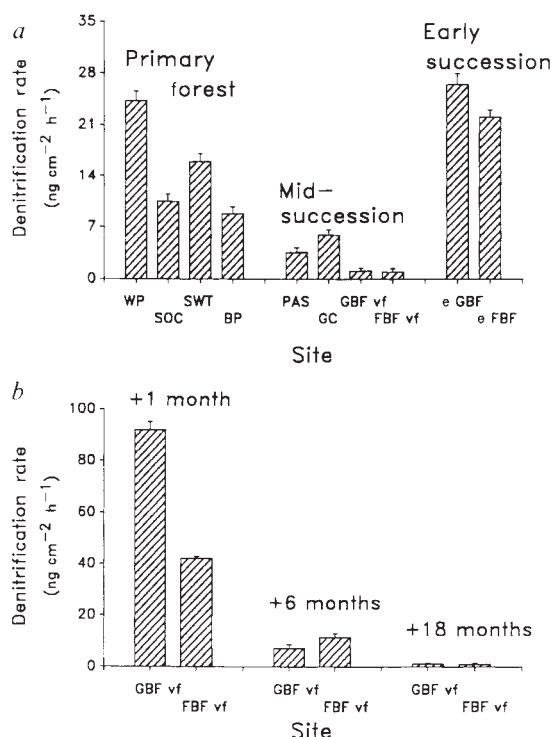


Fig. 2 *a*, Mean denitrification rates ($N_2 + N_2O$ production) in rain-forest sites at different stages of secondary succession (see Fig. 1 legend). Values were log-normal-transformed before analysis to satisfy homogeneity of variance assumptions; vertical lines in bar centres represent 1 standard error. *b*, Denitrification rates in the successional sites described in Fig. 1 legend. Times refer to months following secondary vegetation clearing. vf, vegetation-free; e denotes early.

Despite high within-site variability, several trends were significant. First, denitrification differed significantly ($P < 0.001$) among vegetation groups (primary-forest sites, mid-succession sites and early-succession sites). Rates of denitrification were significantly higher in our primary-forest and early-succession sites than in any of the mid-succession sites ($P < 0.05$) (Fig. 2*a*).

The rates of denitrification shown in Fig. 2*a* extrapolate to a mean of 0.11 g N m^{-2} per month in the primary-forest sites, to 0.16 g N m^{-2} per month in the early-successional sites, and to $< 0.034 \text{ g N m}^{-2}$ per month in the mid-successional sites. These values are within the range of those reported for temperate-region forests^{8,9,18}, although in individual cores denitrification occurred at rates substantially higher than any yet reported. For example, in the primary forest a number of cores denitrified at rates of $300\text{--}900 \text{ ng N cm}^{-2} \text{ h}^{-1}$ (Fig. 1), values that extrapolate to $2.1\text{--}5.4 \text{ g N m}^{-2}$ per month.

Second, very high rates of denitrification in those sites recently cleared of secondary vegetation tapered off to rates close to those in the mid-successional site within a few months after clearing (Fig. 2*b*). In one of the sites, rates averaged $90 \text{ ng N cm}^{-2} \text{ h}^{-1}$ shortly after clearing; within 4 months, mean denitrification rates had fallen to $< 10 \text{ ng N cm}^{-2} \text{ h}^{-1}$. A year later, in the absence of vegetation which might have even further depressed denitrification rates as a result of plant-denitrifier competition for nitrate^{24,25}, rates had fallen to $< 2 \text{ ng N cm}^{-2} \text{ h}^{-1}$, perhaps because of nitrate immobilization during decomposition of grass or fern roots.

Third, in the primary-forest sites, denitrifiers did not respond to added nitrate, whereas nitrate added to soils from mid-successional sites with vegetation stimulated denitrification substantially (Table 1). This suggests that competition for nitrate may be a principal factor in keeping denitrification low in the

Table 1 Response of denitrifiers in intact soil cores from primary-forest and mid-successional sites to added water and water plus nitrate

	$\text{N}_2\text{O}-\text{N}$ ($\text{ng cm}^{-2} \text{ h}^{-1}$)	
Primary forest		
WP	87.7 (5.4)	77.2 (4.4)
SOC	111.0 (5.1)	91.6 (5.2)
BP	12.3 (3.9)	12.7 (0.9)
SWT	49.9 (4.0)	38.1 (9.0)
Mid-succession		
PAS (regrowth)	2.7 (1.5)	9.0 (2.2)*
GC (regrowth)	0.2 (1.4)	148.0 (5.4)*
GBF (vegetation-free)	1.0 (2.5)	0.03 (1.0)
GBF (vegetation-free)	1.1 (1.9)	0.3 (1.3)

Cores were assayed for nitrate response after the initial non-amendment denitrification assay described in the text. Values are means of 10 cores ± 1 standard error; * indicates significant differences at $P < 0.05$.

vegetated (non-bare fallow) mid-successional sites, an explanation consistent with results from temperate-region studies and with general theory concerning factors that regulate denitrification in non-agricultural sites^{25,26}. In other primary-forest sites, such as those in the Amazon, nitrate has been shown to stimulate N_2O fluxes^{27,28}. Presumably in these very infertile soils nitrogen is in short supply even in primary forest.

This study represents the first systematic measurement of denitrification in humid tropical ecosystems. Because of the initial nature of this endeavour, because humid tropical soils are diverse^{29,30}, and because system-wide responses to disturbance such as clear-felling can be dependent on many factors, including the type and severity of disturbance³¹, caution should be exercised in generalizing our results to humid tropical ecosystems worldwide. Before general conclusions can be drawn, we need a broad database covering a variety of tropical ecosystems, and careful consideration of historical patterns of change in clearing rates. Nevertheless, successional theory suggests that similar trends ought to occur in other parts of the humid tropics following similar sorts of disturbance. If this is so, then our results suggest that denitrification may represent an important but transient source of nitrogen loss following tropical-rain-forest clearing, and that deforestation in the past few decades should lead to lower rates of nitrogen gas emission worldwide, as high-activity primary forest is converted to lower-activity mid-successional vegetation.

The importance of these results to our understanding of global N_2O fluxes *per se* is clouded by a generally poor understanding of the biological sources of N_2O in most ecosystems. Denitrifiers—both heterotrophs such as *Pseudomonas* spp. and chemoautotrophs such as *Nitrosomonas* spp.—probably comprise the greatest single source of N_2O in most ecosystems, but evidence for the importance of other biological sources is growing³². Furthermore, we have no evidence that the N_2O portion of the total $N_2 + N_2O$ gas flux evolved by denitrifiers is constant among sites in our ecological sequence. This makes it difficult to infer patterns for denitrifier-derived N_2O fluxes with confidence even for these sites. However, because N_2O and denitrification ($N_2 + N_2O$) fluxes in the systems studied so far appear to be under similar system-level controls^{8,9,18,28,33}, it is conceivable that N_2O fluxes *per se* may be related to plant succession in the same way that denitrification appears to be. Given this likelihood, and the importance of N_2O fluxes to atmospheric processes such as ozone depletion in the stratosphere^{34–36} and infrared transmissivity in the troposphere¹¹, we believe that the relationship between plant succession and both denitrification and N_2O fluxes *per se* deserves clarification in additional tropical and other globally-important ecosystems.

We thank E. Hidalgo and B. Paniagua for technical assistance in Costa Rica. P. M. Groffman, P. Sollins and P. M. Vitousek made many helpful comments on an earlier draft, and M. A.

Huston provided valuable help with site selection. We thank the Organization for Tropical Studies at Duke University and in particular D. A. and D. B. Clark and the staff of the La Selva Biological Station for logistical support. Financial support was provided by the NSF.

Received 11 August; accepted 15 November 1988.

- Natn. Acad. Sci. *Nitrates: An environmental Assessment* (NRC, Washington, DC, 1978).
- Robertson, G. P. *Environment* **28**, 16–21 (1986).
- Bania, A., Lawless, J. C. & Whitter, R. C. *Adv. Space Res. Proc. COSPAR XXV*, 141–153 (Pergamon, New York, 1984).
- Duxbury, J. M., Bouldin, D. R., Terry, R. E. & Tate, R. L. *Nature* **298**, 462–464 (1982).
- Firestone, M. K. *Agronomy Monogr.* **22**, 289–320 (1982).
- Sahrawat, K. L. & Keeney, D. R. *Adv. Soil Sci.* **4**, 103–148 (1986).
- Tiedje, J. M. in *Biology of Anaerobic Microorganisms* (ed. Zehnder, A. J. B.) 179–244 (Wiley, New York, 1988).
- Melillo, J. M., Aber, J. D., Steudler, P. A. & Schimel, J. P. *Ecol. Bull. (Stockholm)* **35**, 217–228 (1983).
- Robertson, G. P. & Tiedje, J. M. *Soil Sci. Soc. Am. J.* **48**, 383–389 (1984).
- Robertson, G. P. & Rosswall, T. *Ecol. Monogr.* **56**, 43–72 (1986).
- Dickinson, R. E. & Cicerone, R. J. *Nature* **319**, 109–114 (1986).
- Rasmussen, R. A. & Khalil, M. A. K. *Science* **232**, 1623–1624 (1986).
- Sancho, F. & R. Mata. *Estudio detallado de suelos: Estacion Biologica La Selva* (Organization for Tropical Studies, Duke Univ., North Carolina, 1988).
- Sollins, P., Sancho, F., Sanford, R. L. & Parker, G. G. in *La Selva: Ecology and Natural History of a Neotropical Rain Forest* (eds McDade, L., Bawa, K., Hespenheide, H. & Hartshorn, G.) (Sinauer Press, Chicago) (in the press).
- Radulovich, R. & Sollins, P. *Soil Sci. Soc. Am. J.* **51**, 1386–1388 (1987).
- Hartshorn, G. S. in *Costa Rican Natural History* (ed. Janzen, D. H.) Ch. 7 (Univ. of Chicago Press, 1983).
- Harcombe, P. A. in *Recovery and Restoration of Damaged Ecosystems* (eds Cairns, J., Dickson, K. L. & Herricks, E. E.) 347–378 (Univ. of Virginia Press, Charlottesville, Virginia, 1977).
- Robertson, G. P., Vitousek, P. M., Matson, P. A. & Tiedje, J. M. *Plant & Soil* **97**, 119–127 (1987).
- Tiedje, J. M. in *Methods of Soil Analysis, Part 2* (eds Miller, R. H. & Keeney, D. R.) 1011–1026 (Am. Soc. Agron., Madison, Wisconsin, 1982).
- Parkin, T. B., Kaspar, H. F., Sextstone, A. J. & Tiedje, J. M. *Soil Biol. Biochem.* **16**, 323–330 (1984).
- Ryden, J. C. & Skinner, J. H. *Soil Biol. Biochem.* **19**, 753–757 (1987).
- Tiedje, J. M., Simkins, S. & Groffman, P. M. *Plant & Soil* (in the press).
- Folorunso, O. A. & Rolston, D. E. *Soil Sci. Soc. Am. J.* **48**, 1214–1219 (1984).
- Smith, M. S. & Tiedje, J. M. *Soil Sci. Soc. Am. J.* **43**, 951–955 (1979).
- Groffman, P. M., Tiedje, J. M., Robertson, G. P. & Christensen, S. *Advances in Nitrogen Cycling in Agricultural Ecosystems* (ed. Wilson, J. R.) 174–192 (Commonwealth Agric. Bur., Sydney, Australia, 1988).
- Robertson, G. P. *J. Ecol.* (in the press).
- Keller, M., Kaplan, W. A., Wofsy, S. C. & DaCosta, J. M. *J. geophys. Res.* **93**, 1600–1604 (1988).
- Livingston, G. P., Vitousek, P. M. & Matson, P. A. *J. geophys. Res.* **93**, 1593–1599 (1988).
- Sanchez, P. *Properties and Management of Soils in the Tropics* (Wiley, New York, 1976).
- Uehara, G. & Gillman, G. *The Mineralogy, Chemistry and Physics of Tropical Soils with Variable Charge Clays* (Westview, Boulder, Colorado, 1981).
- Gorham, E., Vitousek, P. M. & Reiners, W. A. *A. Rev. ecol. Syst.* **10**, 53–88 (1979).
- Robertson, G. P. & Tiedje, J. M. *Soil Biol. Biochem.* **19**, 187–193 (1987).
- Matson, P. A. & Vitousek, P. M. *Global Biogeochemical Cycles* **1**, 163–170 (1987).
- Hahn, J. & Crutzen, P. J. *Phil. Trans. R. Soc. Lond. B* **296**, 521–541 (1982).
- Singh, H. B. *Envir. Sci. Technol.* **21**, 320–327 (1987).
- Cicerone, R. J. *Science* **237**, 35–41 (1987).

Origin of Messel Oil Shale kerogen

K. Goth*, J. W. de Leeuw†, W. Püttmann‡
& E. W. Tegelaar†

* Forschungsinstitut Senckenberg, Senckenberganlage 25,
6000 Frankfurt am Main, FRG

† Delft University of Technology, Faculty of Chemical Technology
and Materials Science, Organic Geochemistry Unit, de Vries van
Heystplantsoen 2, 2628 RZ Delft, The Netherlands

‡ Lehrstuhl für Geologie, Geochemie und Lagerstätten des Erdöls
und der Kohle, RWTH Aachen, Lochnerstrasse 4-20,
5100 Aachen, FRG

Despite many investigations concerning the insoluble organic matter in sediments (kerogen), its chemical nature and origin are only poorly understood. Here we report the results of a combined microscopy and chemical study of the Messel Oil Shale which shed light on the mechanism of kerogen formation. Scanning electron microscopy revealed the overwhelming presence of cell-wall remains of *Tetradron-like* microalgae which are virtually indistinguishable from those of the widely occurring extant *Tetradron minimum* (Chlorococcales). Flash-pyrolysis gas chromatography/mass spectroscopy indicated the presence of an insoluble, non-hydrolysable highly aliphatic biopolymer in both fossil and extant *Tetradron* species. The bulk of the Messel Oil Shale kerogen probably consists of selectively preserved cell-wall material of *Tetradron* algae, mainly made up of this newly discovered biopolymer. We therefore suggest that this polymer, and similar types of recently discovered highly aliphatic biopolymers in other algae and plant cuticles, are important precursors of *n*-alkanes in crude oils.

The Messel Oil Shale (near Darmstadt, West Germany) represents an organic-matter-rich lacustrine deposit of Mid-Eocene (~48-Myr) age¹. This sediment (in particular, its fossil content) has been studied extensively by palaeontologists and organic geochemists², but despite these efforts, the origin and nature of the bulk of the organic matter is poorly understood.

Microscopic examinations of a large number of samples, obtained from outcrops within the open-cast mine at the aforementioned location as well as samples from cores, revealed that the organic matter in the Messel Oil Shale (~30% on a dry-weight basis) is mainly concentrated in distinct laminae. These organic-rich laminae alternate with laminae that are rich

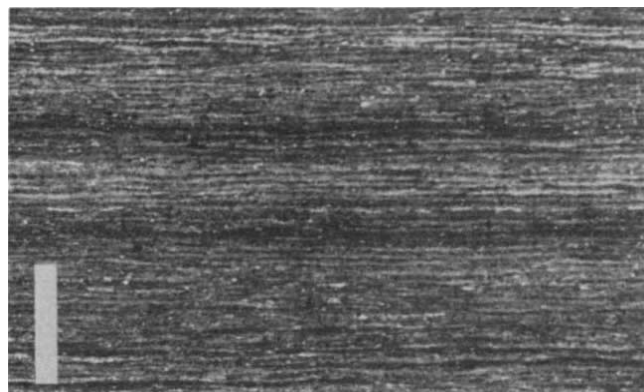


Fig. 1 Cross-section of a typical laminated oil shale from the Messel Formation. Light-coloured microlaminae are composed of *Tetradron* remains, and dark layers represent background sediment. Scale bar, 2.0 mm.

in clay and which also contain minor amounts of terrestrially derived organic debris (Fig. 1). Such units of microlaminae are commonly ~0.1–0.2 mm thick and their abundant occurrence throughout the Messel Oil Shale sequence suggests a cyclic sedimentation pattern (K.G., manuscript in preparation). The lack of vertical circulation within the lake, probably caused by density stratification, prevented bioturbation, thus permitting the preservation of these laminae. The occurrence of pyrite and siderite^{3,4} supports the suggestion of an oxygen-depleted deep water body. Further evidence for density stratification is provided by the presence of porphyrins related to the bacteriochlorophyll *d* series, which originate from the anaerobic phototrophic bacteria *Chlorobiaceae*⁵.

Examination of a large number of the organic-rich laminae by scanning electron microscopy (SEM) revealed a predominance of small sculptured bodies, relatively uniform in size (5–10 µm, Fig. 2), which are considered to be the cell-wall remains of unicellular algae⁴. A sharp delineation was observed between the base of these nearly pure algal laminae and the inorganic laminae. This alternating sequence results from the sedimentation of algal debris, originating in annual blooms of these organisms, superimposed on the mainly inorganic, terrestrially derived background (K.G., manuscript in preparation).

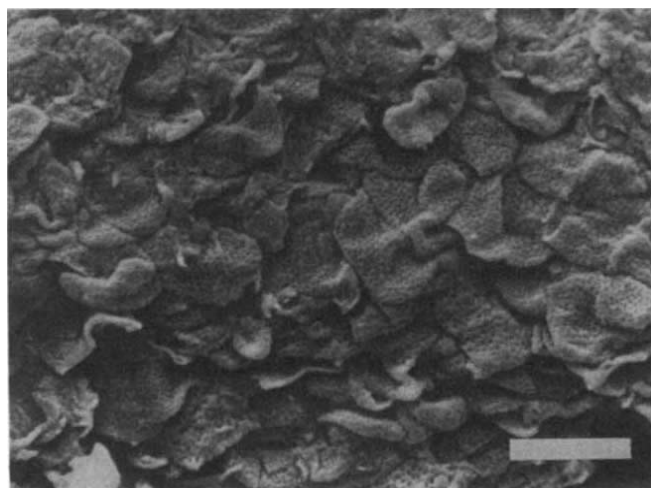


Fig. 2 Scanning electron micrograph image of a *Tetraedron* micro-lamina from the Messel Oil Shale. Note the dense packing of cell-wall remains. Scale bar, 10 μm .

The cell-wall remains comprise the most common biogenetic component and are thus believed to represent the predominant part of the organic matter in the Messel Oil Shale (K.G., manuscript in preparation). Their accumulation in distinct layers made it possible to obtain almost pure samples of algal-derived organic matter by physically removing the algal laminae from the background sediment using a sharp scalpel. The algal cells are pillow-shaped, with either a wart- or net-like surface structure (Fig. 3a)^{4,6}. Transmission electron microscopy (TEM) studies of cross-sections revealed internally homogeneous wavy cell walls varying in thickness from 0.04 μm between, and 0.2 μm at, the corrugations. All these features (size, shape, surface pattern, preserved wall structure) are very similar to a species of the extant chlorococcal genus *Tetraedron* (K.G., manuscript in preparation).

The outer morphology of the Messel fossils is virtually indistinguishable from the extant *Tetraedron minimum* (Fig. 3b). Its single cells are quadrangular and flattened (pillow-like), 5–20 μm long and 3–8 μm thick. The angles are rounded, tapering into a short papilla^{7,8} with surfaces covered either by small verrucae or by an irregular network, depending on the cell size. The thick cell wall is composed of three layers⁹. This cosmopolitan species occurs abundantly in the plankton of slightly eutrophized fish-ponds, lakes and rivers^{7,8}.

Tetraedron cell-wall remains appear to be by far the most important constituents of the organic matter in the Messel Oil Shale. We therefore attempted to chemically characterize the insoluble cell-wall material of the extant *T. minimum* and its fossil counterpart using flash-pyrolysis gas chromatogra-

phy/mass spectrometry (Py-GCMS), to shed light on the formation and structure of the Messel shale kerogen. It must be noted that the assumed contribution of several other types of microbial organisms, such as dinoflagellates and chlorobium-type bacteria, to the organic matter of the Messel Oil Shale is based mainly on lipids, which comprise at most 5% of the organic matter^{5,10}.

The Py-GCMS total-ion chromatogram of a Messel Oil Shale sample, highly enriched with *Tetraedron* cell-wall remains, and that of the isolated cell walls of *T. minimum*, are shown in Fig. 4. Both chromatograms are dominated by homologous series of straight-chain α , ω -alkadienes, alk-1-enes and alkanes ranging from C_6 to C_{31} . The longer-chain alkanes show a slight preference for odd, rather than even, carbon number. The strong similarities between both pyrolysates indicates the presence of an insoluble, non-saponifiable aliphatic macromolecular structure in the cell walls of the extant *T. minimum* and its fossil counterpart. Based on these results, the *Tetraedron* cell-wall polymer appears to be similar to the insoluble, non-hydrolysable highly aliphatic biopolymers present in plant cuticles^{11–13} and all three races of *Botryococcus braunii* algae^{14,15}. It should be noted, however, that in the extant *T. minimum* this macromolecular structure comprises only a minor part (~5 wt%; E.W.T., unpublished results) of the total biomass, whereas it is abundantly present in the *Tetraedron*-enriched Messel sample. During diagenesis, the labile components (for example, polysaccharides) of the *Tetraedron* cells are presumably degraded, leading to relative enrichment of the aliphatic cell-wall biopolymer in the Messel Oil Shale kerogen. The selective preservation and relative enrichment of similar biopolymers has been observed previously (refs 14, 15, and E.W.T., J. H. F. Kerp and J.W. de L., manuscript in preparation).

On the basis of the accumulating evidence for the direct contribution of *in situ* insoluble organic substances to kerogen^{11–18} and of the abundant presence of *n*-alkanes and *n*-alkenes in the pyrolysates of almost any oil-prone kerogen¹⁹, it is suggested that the selective preservation of resistant macromolecular structures represents an alternative mechanism for the formation of kerogen (E.W.T., J. H. F. Kerp and J.W. de L., manuscript in preparation). This contrasts with the currently accepted theory whereby kerogen results from an abiological random polymerization of monomers derived from the biodegradation of biopolymers, such as polysaccharides and proteins, and from existing aliphatic lipid moieties^{20,21}.

Kerogens with a suitable hydrogen/carbon ratio are believed to generate oil at high temperature and pressure²⁰. It has been speculated that *n*-alkanes present in crude oils are thermally derived from the aliphatic lipid moieties through defunctionalization²⁰. Maturation experiments have already shown, however, that the insoluble, non-hydrolysable highly aliphatic biopolymer, isolated from the cuticular membranes of the extant *Agave americana* (E.W.T. *et al.*, manuscript in preparation), yield long-chain alkanes under thermal stress, indicating the oil potential of these aliphatic biopolymers.

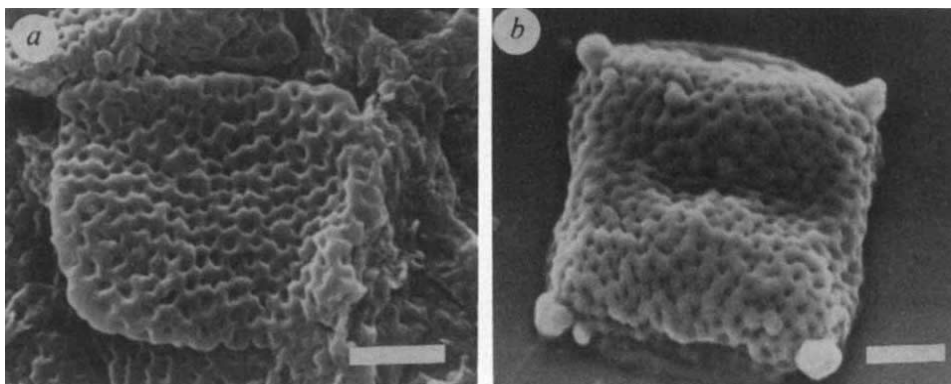
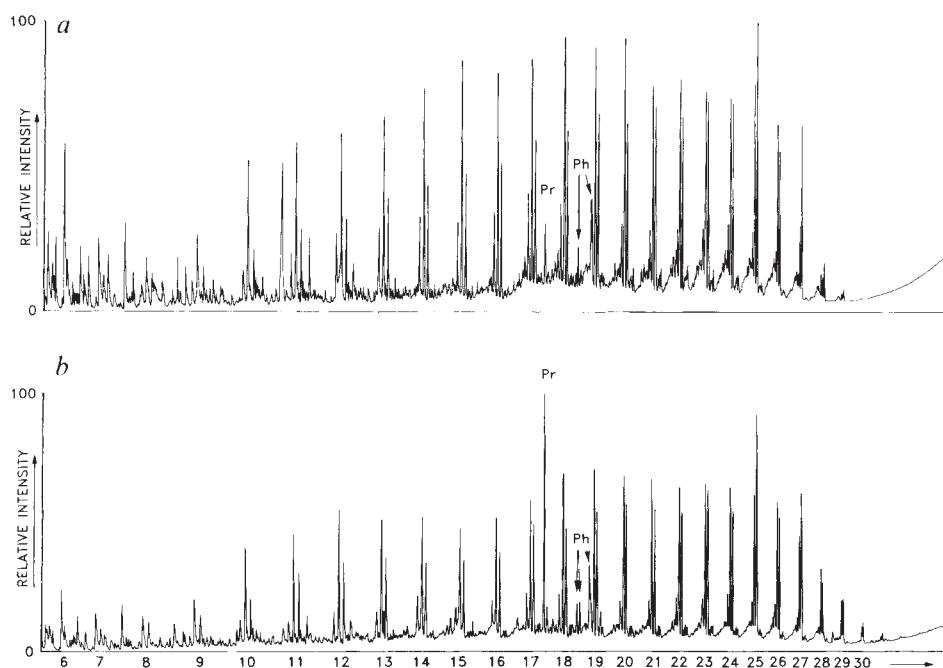


Fig. 3 Scanning electron photomicrographs of single *Tetraedron* cells. *a*, *Tetraedron* sp. from the Eocene Messel Oil Shale; scale bar, 2 μm . *b*, Extant *T. minimum* (strain SAG 44.81); scale bar, 2 μm . Notice the similarity in the surface patterns. Papillar processes at the cell angles do not occur or are not preserved in the Messel *Tetraedron* sp.

Fig. 4 Total-ion chromatograms of Py-GCMS analyses of *a*, isolated cell wall of *T. minimum*, and *b*, *Tetraedron*-enriched Messel Oil Shale. Isolation of the cell walls of *T. minimum* was achieved by extraction, saponification and acid hydrolysis, carried out according to methods described previously (refs 22, 23 and M. E. C. Moers *et al.*, manuscript in preparation). Details of the Py-GCMS analyses have been described before²⁴. Instrumental conditions were: Curie-point temperature 770 °C; Gas-chromatographic separation was achieved on a fused silica column (length = 25 m; inner diameter = 0.31 mm) coated with chemically bonded CPSil 5; temperature programme of the GC oven comprised heating from 0 to 320 °C at 3 °C min⁻¹, then holding at 320 °C for 20 min; mass spectra were obtained at an ionization voltage of 70 eV. Peak identification is as follows: the numbers refer to the homologous series of α, ω -alkadienes, *n*-alk-1-enes and *n*-alkanes; Pr denotes pristenes, probably originating from bound tocopherol²⁵; Ph denotes phytadienes, probably originating from chlorophylls²⁶.



Several studies are currently being performed to elucidate the structures of these types of resistant aliphatic biopolymers and to determine the extent of their occurrence.

The unusually good preservation of organisms in the Messel Oil Shale thus enables scientists to investigate interrelationships of fundamental importance in geosciences. The plans for rubbish disposal at the former open-cut mine in Messel should therefore be abandoned to prevent a unique site from being lost forever.

The unialgal culture for the SEM and geochemical studies was obtained from the collection of algae of the Institute for Plant Physiology, University of Göttingen (strain SAG 44.81, *T. minimum*). We would like to thank Dr Mollenhauer of the Algological Department of the Research Institute, Senckenberg, who was engaged in the cultivation of *Tetraedron* strains, and M. Baas of the TU Delft for technical assistance. This work was partly supported by the Deutsche Forschungsgemeinschaft (DFG) (Projekt Schaarschmidt 178/6) and the Netherlands Foundation for Earth Science Research (AWON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

Received 25 July; accepted 3 November 1988.

1. Tobien, H. *Notizbl. hess. Landesamt. Bodenforsch.* **96**, 111–119 (1968).
2. Von Koenigswald, W. & Michaelis, W. *Geol. Jb. Hessen* **112**, 5–26 (1984).
3. Irion, G. *Natur Mus., Frankf.* **107**, 213–218 (1977).
4. Goth, K. *Cour. Forsch.-Inst. Senckenberg* **85**, 209–211 (1986).
5. Hayes, J. M., Takigiku, R., Ocampo, R., Callot, H. J. & Albrecht, P. *Nature* **329**, 48–51 (1987).
6. Püttmann, W. & Goth, K. *Cour. Forsch.-Inst. Senckenberg* **102** (in the press).
7. Kováčik, L. *Arch. Hydrobiol. Suppl.* **46**, 354–391 (1975).
8. Komárek, J. & Fott, B. *Das Phytoplankton des Süßwassers, 7. Teil, 1. Hälfte, Chlorophyceae (Grünalgen) Ordnung: Chlorococcales 695–700* (Schweizerbart, Stuttgart, 1983).
9. Kováčik, L. & Kalina, T. *Arch. Hydrobiol. Suppl.* **46**, 433–444 (1975).
10. Ourisson, G., Mattern, G. & Albrecht, P. *J. chem. Soc. D* **22**, 1570–1571 (1970).
11. Nip, M., Tegelaar, E. W., de Leeuw, J. W., Schenck, P. A. & Holloway, P. J. *Naturwissenschaften* **73**, 579–585 (1986).
12. Nip, M. *et al.* in *Advanced Organic Geochemistry 1985* (eds Leythaeuser, D. & Rullkötter, J.) 769–778 (Pergamon, Oxford, 1986).
13. Nip, M. *et al.* *Geochim. cosmochim. Acta* (in the press).
14. Largeau, C., Derenne, S., Casadevall, E., Kadouri, A. & Sellier, N. in *Advanced Organic Geochemistry 1985* (eds Leythaeuser, D. & Rullkötter, J.) 1023–1032 (Pergamon, Oxford, 1986).
15. Derenne, S., Largeau, C., Casadevall, E., Tegelaar, E. W. & de Leeuw, J. W. *Fuel Process. Tech.* (in the press).
16. Chappe, B., Albrecht, P. & Michaelis, W. *Science* **217**, 65–66 (1982).

17. Philp, R. P. & Calvin, M. *Advanced Organic Geochemistry 1975* (eds Campos, R. & Goni, J.) 735–752 (Enadisma, Madrid, 1976).
18. Philp, R. P. & Calvin, M. *Nature* **262**, 134–136 (1976).
19. Van de Meent, D., Brown, S. C., Philp, R. P. & Simoneit, B. R. T. *Geochim. cosmochim. Acta* **44**, 999–1013 (1980).
20. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence*. (Springer, Berlin, 1984).
21. *Kerogen—Insoluble Organic Matter from Sedimentary Rocks* (ed. Durand, B.) (Technip, Paris, 1980).
22. Holloway, P. J. in *CRC Handbook of Chromatography, Lipids Vol. 1* (ed. Mangold, H. K.) 321–334 (CRC, Boca Ranton, 1984).
23. Holloway, P. J. in *CRC Handbook of Chromatography, Lipids Vol. 1* (ed. Mangold, H. K.) 347–380 (CRC, Boca Ranton, 1984).
24. Tegelaar, E. W. *et al.* *J. anal. appl. Pyrolysis* (in the press).
25. Goossens, H., de Leeuw, J. W., Schenck, P. A. & Brassell, S. C. *Nature* **312**, 440–442 (1984).
26. Van de Meent, D., de Leeuw, J. W. & Schenck, P. A. *Adv. org. Geochem. 1979* (eds Douglas, A. G. & Maxwell, J. R.) 469–474 (Pergamon, Oxford, 1980).

The critical slip distance for seismic faulting

C. H. Scholz

Lamont-Doherty Geological Observatory and Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA

Experimentally based friction laws^{1,2} have been found to predict virtually the entire range of observed behaviour of natural faults^{3,4}. These laws contain a critical slip distance, L , which plays a key role in determining the degree of fault instability, the size of the zone of earthquake nucleation, the frictional breakdown width, and the proportion of pre- and post-seismic slip to co-seismic slip. In laboratory measurements L is found to be about 10^{-5} m, but modelling results show that it must be about 10^{-2} m if natural earthquake behaviour is to be simulated. The discovery that fault surfaces are fractal over the scale range 10^{-5} – 10^5 (refs 5, 6), even for faults with large net slip, has confused the problem of scaling this parameter from laboratory experiments to natural faults, because fractal surfaces have no characteristic length. Here I show that geometrically unmated fractal surfaces, when in contact under

a normal load, develop a characteristic length in their contact because long-wavelength apertures close under load whereas short-wavelength apertures may remain open. This critical distance may be identified with L , and calculations based on fault topography data show that at seismogenic depths it will be in the range anticipated from the earthquake modelling studies.

The critical slip distance L is the distance that slip must occur for sliding surfaces, upon a slip velocity perturbation, to change from one steady-state friction condition to another. It is interpreted as being the distance required to change entirely the population of asperity contacts¹, and so is related to the geometry of contact. It has been shown that friction becomes completely uncorrelated after slipping a distance equal to twice the mean contact junction diameter, D_j , and falls to an autocorrelation value of $1/e$ after slip of about one junction diameter⁷. L is defined as the slip distance for friction to evolve to $1/e$ from its instantaneous value following a velocity jump, and so should equal D_j . L for metallic friction has been found to be in good agreement with estimates of D_j determined from autocorrelation analysis and direct measurement⁷. Laboratory data on rock friction indicate that L increases with surface roughness⁸, which is consistent with this model. But surfaces used in laboratory friction experiments are artificially ground and have a corner frequency in their power spectra, and hence possess a maximum contact spacing and minimum mated junction diameter, whereas natural surfaces are fractal and have no such characteristic length⁵. This produces a problem in estimating L for natural surfaces from the basic friction data established in the laboratory. If no characteristic distance exists for faults, the application of such friction laws to earthquake faulting must be questioned.

A thought experiment is helpful here. Fractal surfaces have asperity amplitudes that increase with wavelength. A fault consists of two such surfaces in contact, which must be geometrically mismatched at all distances less than the net fault slip. But no macroscopic aperture may be observed between natural fault surfaces, and this cannot be entirely explained as the result of filling by such fluid-like materials as clay fault gouge (in solid friction we may initially ignore such weak materials and consider that a gouge-filled aperture is the same as an empty aperture). The fault must therefore have closed, but little evidence is found of inelastic flow that would allow such closure, and the physical conditions that would allow such inelastic deformation are not always present. Although some closure may occur by slip on systems of secondary faults, in the following we treat the closure as elastic.

Consider the contact between two surfaces to be composed of a network of bridges, of varying length λ , between asperities of varying height h . We approximate the behaviour of the bridges by treating them as cracks, with aspect ratio h/λ . An order-of-magnitude approximation for the pressure, p , required to completely close a crack elastically is⁹

$$p \approx E \frac{h}{\lambda} \quad (1)$$

where E is Young's modulus. For fractal surfaces h and λ are related. For the special case of Brownian surfaces,

$$h^2 = \kappa \lambda \quad (2)$$

where κ is a constant. This result may be generalized for fractals of other dimension, but such is not required here. Combining equations (1) and (2), we obtain the following expression for the maximum length of bridge that may remain open under a pressure p :

$$\lambda_c = \left(\frac{E}{p} \right)^2 \kappa \quad (3)$$

This result shows that Brownian surfaces in contact under a pressure p will be mated at all wavelengths longer than λ_c . The mechanical behaviour of contacting surfaces depends only on

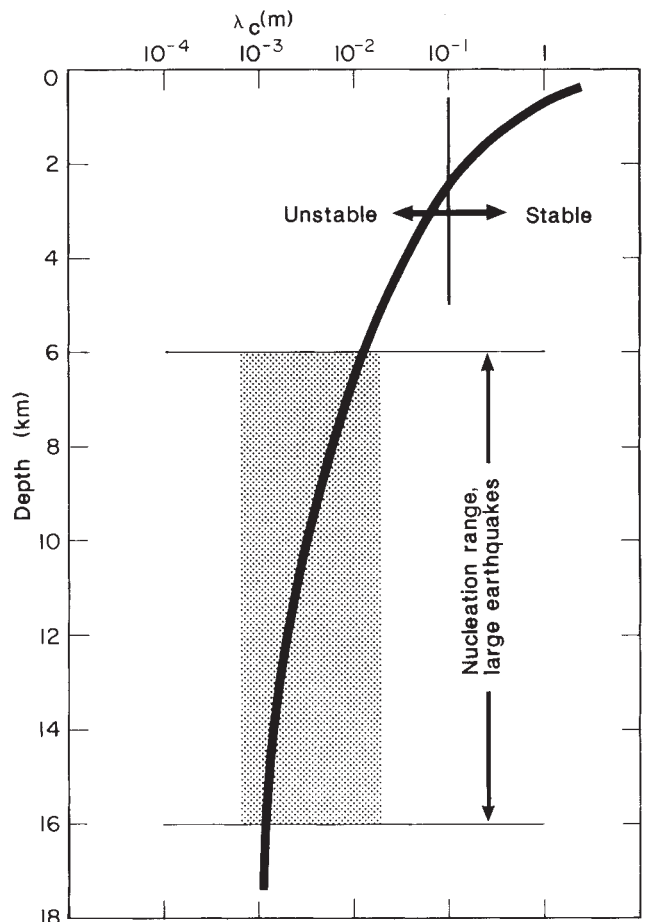


Fig. 1 The minimum mated contact junction diameter, λ_c , calculated as a function of depth in the Earth. This parameter may be interpreted as equal to L , the critical slip distance. Effective pressure is assumed to increase at a rate of 15 MPa km^{-1} .

the unmated part of the contact spectrum¹⁰, so that, in terms of contact properties, λ_c appears as a characteristic length.

In particular, λ_c is the maximum contact spacing and it is also the minimum diameter of mated junctions. Because the zero-set at an arbitrary level of a fractal function has a power-law interval distribution¹¹, the size distribution of junctions will also obey a power law, down to this minimum size. This indicates that there will be a spectrum of critical slip distances, as has been suggested to be manifest in a more general form of the rate- and state-variable friction laws². We discuss only the shortest of these, $L \approx D_{j,\min} \approx \lambda_c$, because that will be what determines slip stability, and because, as a result of the power-law size distribution, this junction diameter will be the most common and will be expected to dominate the behaviour.

Equation (3) can be evaluated with the use of fault topography data, which have been taken in the direction parallel to slip and presented as plots of power spectral density¹². The power-spectral density, G , of a Brownian surface is $G(1/\lambda) = \kappa \lambda^2$. Although the data used do not have a constant spectral slope, this equation fits them reasonably well over the range 5×10^{-4} – 5×10^{-1} m, which is the range relevant to these calculations. These data yield $\kappa = 2 \times 10^{-8} \text{ m}$, and using the representative value $E = 7 \times 10^4 \text{ MPa}$, λ_c may be calculated as a function of depth, as shown in Fig. 1.

This result shows that λ_c decreases with depth. Over the depth range in which large earthquakes nucleate (6–16 km), it lies in the range 10^{-2} – 10^{-3} m. If, as we have assumed, $L = \lambda_c$, then this result is in agreement with values of L required by modelling studies to simulate geophysically observed behaviour^{3,4}. The modelling studies indicate that if L is too large, the fault will

be stable and will not generate earthquakes. The value of L at the stability transition, L_s , depends on the assumed value of $(a-b)$, the parameter defining the slip velocity dependence of friction⁴, and is $\sim 10^{-1}$ m (see Fig. 1). The result indicates that a stability transition will occur at a depth of ~ 2.5 km, which is additional to that expected from the action of fault gouge in changing the sign of $(a-b)$ ¹³.

This effect may be estimated in a different way, from an expression for the maximum size of a stable fault patch, which is also a function of L and p . Combining equation (3) with equation (14) of ref. 14, we obtain

$$r_c = \frac{0.37E^3\kappa}{p^3(b-a)} \quad (4)$$

The transition will occur at a depth equal to this maximum stable dimension. Assuming a representative value of $(b-a) = 0.005$, equation (4) indicates that this depth is ~ 3.5 km, in reasonable agreement with our first estimate.

Previous modelling studies have assumed a constant value of L . Here we show that it must decrease with depth. This will result in a degree of fault instability that increases with depth, which will enhance the tendency for large earthquakes to nucleate near the base of the seismogenic region¹⁵. An earthquake nucleation model¹⁴ based on this type of friction law indicates that during nucleation there will be a precursory

slip, $u_n \approx 5L$, on a patch of radius $r_n \approx 10^5 L$. From the values of L estimated here, $5 > u_n > 0.5$ cm and $1 > r_n > 0.1$ km, with the smaller values obtaining at greater depth, where large earthquakes tend to nucleate. This will make more difficult the problem of detecting, with surface measurements, precursory phenomena resulting from nucleation.

This work was supported by the US Geological Survey. I thank N. Yoshioka for early discussions of this problem, and F. Chester, K. Evans, A. Lerner-Lam and J. Dieterich for helpful reviews.

Received 16 August; accepted 7 November 1988.

1. Dieterich, J. H. *J. geophys. Res.* **84**, 2161–2168 (1979).
2. Ruina, A. L. *J. geophys. Res.* **88**, 10359–10370 (1983).
3. Tse, S. T. & Rice, J. R. *J. geophys. Res.* **91**, 9452–9472 (1986).
4. Stuart, W. D. *PAGEOPH* **126**, 619–641 (1988).
5. Brown, S. R. & Scholz, C. H. *J. geophys. Res.* **90**, 12575–12582 (1985).
6. Scholz, C. H. & Aviles, C. A. in *Earthquake Source Mechanics* (eds Das, S., Boatwright, J. & Scholz, C.) 147–155, (Am. Geophys. Un. Washington, DC, 1986).
7. Rabinowicz, E. *J. appl. Phys.* **22**, 1373–1379 (1951); *J. appl. Phys.* **27**, 131–135 (1956); *Proc. phys. Soc. Lond.* **71**, 668–675 (1958).
8. Okubo, P. G. & Dieterich, J. H. *J. geophys. Res.* **89**, 5817–5827 (1984).
9. Walsh, J. B. *J. geophys. Res.* **70**, 381–389 (1965).
10. Brown, S. R. & Scholz, C. H. *J. geophys. Res.* **90**, 5531–5545 (1985).
11. Mandelbrot, B. B. *The Fractal Geometry of Nature* 236 (Freeman, San Francisco, 1983).
12. Power, W. L., Tullis, T. E. & Weeks, J. D. *J. geophys. Res.* **93** (in the press).
13. Marone, C. & Scholz, C. H. *Geophys. Res. Lett.* **15**, 621–624 (1988).
14. Dieterich, J. H. in *Earthquake Source Mechanics* (eds Das, S., Boatwright, J. & Scholz, C.) 37–47 (Am. Geophys. Un. Washington, DC, 1986).
15. Das, S. & Scholz, C. H. *Nature* **305**, 621–623 (1983).

Pre-Incan agricultural activity recorded in dust layers in two tropical ice cores

L. G. Thompson, M. E. Davis,
E. Mosley-Thompson & K.-b. Liu*

Byrd Polar Research Center, 125 South Oval Mall,
Ohio State University, Columbus, Ohio 43210, USA

* Department of Geography and Anthropology,
Louisiana State University, Baton Rouge, Louisiana 70803, USA

In South America, no documentary records survive from before the incursion of the Spaniards in the mid-sixteenth century. Knowledge of pre-Hispanic civilizations in this region comes largely from archaeological excavations, which may now be supplemented by information recorded in two ice cores drilled in the Quelccaya ice cap in southern Peru (Fig. 1)^{1,2}. A 1,500-year record of particle concentrations and conductivity in these cores shows two major dust episodes, each lasting about 130 years, centred at AD 920 and 600. Here we examine these dust events in the light of oxygen isotope and net accumulation records, and suggest a correlation with pre-Incan agricultural activity. The annual net accumulation record for AD 1915–1984 compares well with annual changes in Lake Titicaca water levels and with annual precipitation at El Alto (LaPaz, Bolivia), suggesting that the 1,500-year net accumulation record from Quelccaya may serve as a proxy for water level changes in Lake Titicaca. The ice-core record suggests that climatic variability, reflected in lake level changes, has strongly influenced fluctuations in agricultural activity in southern Peru.

The ice-core timescales were reconstructed from a combination of stratigraphic, microparticle concentration (MPC), liquid conductivity (C), and oxygen isotopic ($\delta^{18}\text{O}$) data². As the annual layers thin toward the bottom, the timescales become less precise¹. The records derived from the Quelccaya ice cores show naturally occurring events—marked by increased MPC—such as droughts, the Little Ice Age and volcanic eruptions^{1,2}. Here we concentrate on two unusually intense dust peaks occurring before AD 1000 and explore a possible relationship to pre-Incan cultural activity.

Figure 2 illustrates the decadal averages from AD 470 to 1980 for total MPC, C, $\delta^{18}\text{O}$ and net accumulation (metres of ice-

equivalent). From AD 1000 to 1980 the decadal averages for both cores are combined to produce a composite record. From AD 470 to 1000 only core 1 is used, as the summit core contains a discontinuity at 153.68 m (corresponding to AD 745)¹. The most significant climatic event in the past 1500 years is the Little Ice Age, dated here as AD 1500–1900². Nevertheless, the dust events of the sixth and tenth centuries, recorded by insoluble (MPC) and soluble (C) particles dominate all other features (Fig. 2). These peaks exhibit the greatest intensities around AD 920 and 600 and each lasts ~ 130 yr. The tenth century dust event is equally prominent in both cores.

Little evidence is available to determine the cause of these events, as no written records exist from pre-Hispanic times.

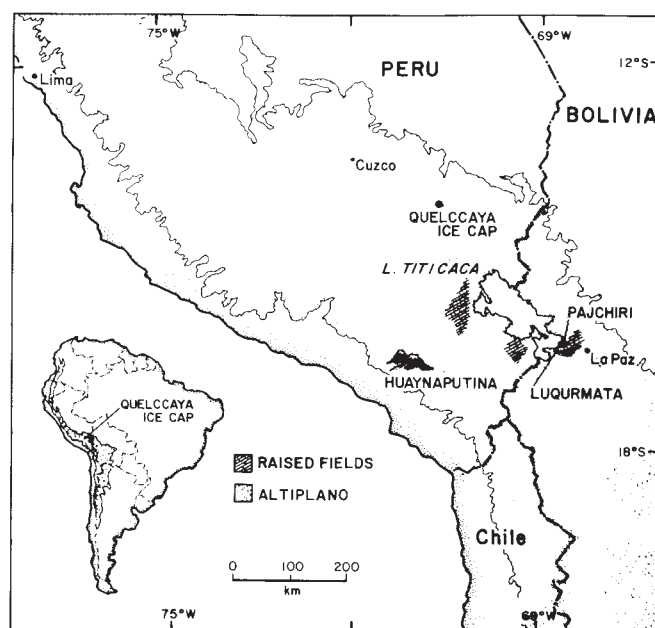


Fig. 1 Map showing the Quelccaya ice cap on the altiplano of Southern Peru ($13^{\circ}56'S$, $70^{\circ}50'W$, 5,670 m elevation) relative to Lake Titicaca and the raised fields.

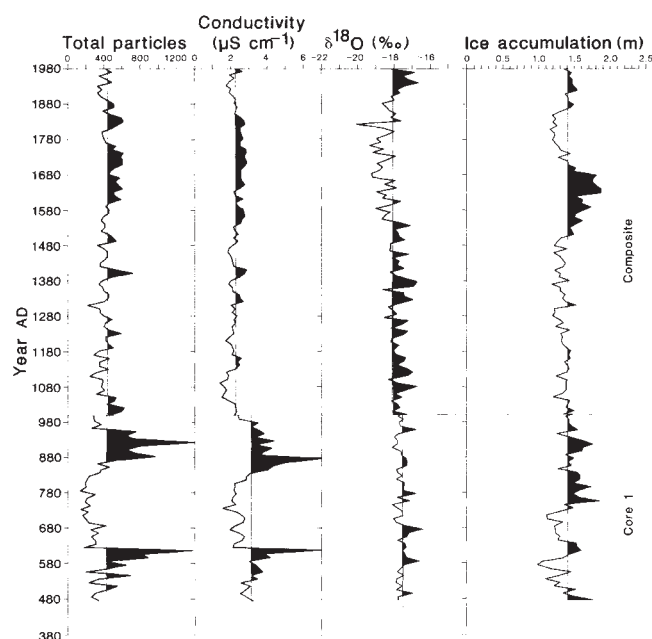


Fig. 2 Profiles of decadal averages from AD 470 to 1980 of total particles (≥ 0.63 to $\leq 16.0 \mu\text{m}$ in diameter per ml of sample $\times 10^3$), conductivity, $\delta^{18}\text{O}$ and net accumulation (metres of ice equivalent). The dust events of AD 920 and 600 are dominant in the microparticle and conductivity profiles.

Therefore, we investigated first the natural causes of perturbations, as characterized by increased MPC and C, observed in the more recent portions of the Quelccaya records.

Water samples from the two dust peaks and from intervening sections in the pre-AD 1000 section of core 1 were examined by light microscopy and scanning electron microscopy (SEM). The particles appear to derive from wind-blown dust, with minor amounts of diatoms and volcanic particles also present. Volcanic activity, however, was not the primary source for these events. Although several active volcanoes are located in southern Peru, the dominant easterly flow inhibits deposition on Quelccaya. Only the eruption of Huaynaputina in AD 1600 (Fig. 1) is distinctly observed in the MPC and C records of the ice cores².

In more recent portions of the core, prolonged droughts produced high soluble- and insoluble-particle concentrations³. But the accumulation profiles demonstrate, when examined on an annual basis, that the peak influx of dust for both events (AD 600 and AD 920) coincided with higher-than-average accumulation. Similarly, the corresponding oxygen isotope record shows no evidence of brief cold episodes during those periods. This behaviour is not characteristic of the neoglaciation period—the 'Little Ice Age', AD 1490–1900 (Fig. 2)—during which the increase in dust was associated with the cooler temperatures inferred from lower $\delta^{18}\text{O}$ (ref. 2).

These two earlier dust events do not appear to be attributable directly to natural sources (such as drought or volcanic activity),

and we therefore consider the possibility of cultural impact on the environment as a major factor in the production of these events.

The Lake Titicaca region, ~200 km south-east of Quelccaya, was once a major centre of pre-Incan cultural and agricultural activity. The altiplano is a cold, wind-swept environment subject to a marked alternation between dry (April–October) and wet (November–March) seasons. The high plateau ranges in elevation from 3,500 m to over 4,000 m and exhibits substantial diurnal temperature variation. During the growing season, frosts and sporadic hailstorms may cause extensive crop damage⁴. The topography, elevation and attendant cold climate of the altiplano severely constrain agriculture in the Titicaca basin. For example, only hardy tubers such as potato, oca, olluco and mashwa, and the unique cold-adapted chenopod grains, quinoa and caníwa, can be readily cultivated. These crops, cultivated extensively in pre-Hispanic times, continue to play a prominent role in the diet of the modern inhabitants of the altiplano^{5,6}.

Around the edges of Lake Titicaca, ridged fields with distinctive patterns have been observed since 1966 (Fig. 1). In pre-Hispanic time, they covered at least 202,800 acres over the Western lake plain at elevations of 3,800–3,850 m (ref. 7) making them the most extensive of their kind in South America⁸. Studies^{9,10} of human activities at Luqurmata and Pajchiri sites on the southern end of Lake Titicaca demonstrate that the peak of raised field construction and maintenance of Tiwanaku V occurred mainly between AD 400 and 900, and that most of the fields were abandoned from AD 900 to 1100. There is no consistent archeological picture, however, of when other sites around the Lake Titicaca basin were occupied. The finer particle sizes in the ice cores from the AD 920 and 600 dust events suggest a lake sediment source.

Under current conditions crops are grown in the rainy season (November–April), when the easterly winds dominate. The crops are harvested during the dry season and the winds come from the south to north-west off the dry altiplano when the fields are bare. The annual dust layers on Quelccaya currently form during the dry season¹¹. Water levels in Lake Titicaca, measured monthly from 1914 to present, show a low water level during each year's dry season¹². In the past, the combination of extensive use of fields and the exposure of large areas of silty sediment as a result of low water levels in regional lakes (including Lake Titicaca) could have produced an important dust source in this region. Located upwind of the ice cap, these fields would increase dust deposition in the snowfall.

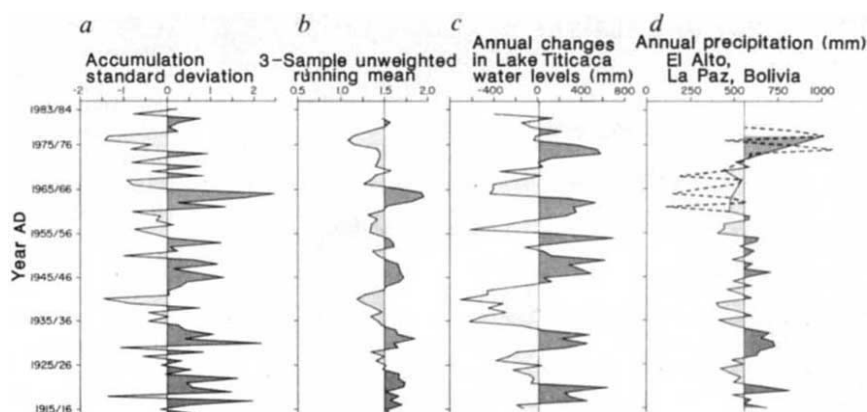
The size distributions of these dust peaks were measured in 14 ranges from 0.63 to 16.0 μm in diameter. The volume per cent distributions are bimodal, with peaks in both fine (0.63–1.6 μm) and coarse (8–16 μm) fractions, but with the former dominating. In contrast, the distributions in most other sections of the core are generally skewed towards the coarser sizes. SEM and light microscopy observations confirm these analyses. The finer particle sizes in the ice cores from the AD 920 and AD 600 dust events therefore suggest a lake sediment source. Furthermore, during expanded use of agricultural fields, pastures would be forced to higher elevations and may have served as dust sources when overgrazing occurred.

Table 1 Coefficients of determination (R^2) for the AD 1915–1961 records of net accumulation from Quelccaya, changes in Lake Titicaca water levels, and precipitation measured at El Alto, Bolivia

	Net accumulation (3-yr r.m.)*	Net accumulation†	Lake T‡	Lake T(+1)§	Lake T(+2)
El Alto¶	0.124	0.116	0.325	0.235	0.022
Net accumulation (3-yr r.m.)*		0.172	0.302	0.324	0.230
Net accumulation†			0.160	0.138	0.014

* Three-year running mean. † Quelccaya net accumulation (standard deviations). ‡ Lake Titicaca annual water level (mm) changes. § Lake T with one-yr lag. || Lake T with 2-yr lag. ¶ Annual precipitation (mm), El Alto (La Paz, Bolivia).

Fig. 3 Quelccaya accumulation record (mm ice equivalent) for AD 1915–1984, presented as standard deviations of annual values (a) and the three-year running mean (b), is compared with annual changes in Lake Titicaca water levels (mm) (c) and annual precipitation (mm) (d) at El Alto (La Paz, Bolivia). Dashed line in (d) represents years with incomplete data.



The Quelccaya accumulation record for AD 1915–1984 is compared with the annual changes in Lake Titicaca water level in Fig. 3. The coefficient of determination ($R^2 = 0.32$) between annual precipitation at El Alto (La Paz, Bolivia) and Lake Titicaca level changes is similar to that ($R^2 = 0.30$) between the three-year running mean of accumulation from Quelccaya and the Lake Titicaca level changes. The coefficients of determination (Table 1) are highest between annual changes in Lake Titicaca water levels (for both the current year and a one-year lag), the Quelccaya three-year average of net accumulation, and annual precipitation at El Alto. Thus, for the period of comparison, the accumulation record from Quelccaya ice cores provides a reasonably accurate record of regional precipitation, suggesting that the 1500-yr Quelccaya net balance record can be used as a proxy for lake-level changes in Lake Titicaca. The relatively low correlations between changes in the annual water levels in Lake Titicaca, the measured precipitation at El Alto, and the Quelccaya net balance record probably result because the lake level is determined not only by regional precipitation but also by evaporation, surface discharge from the lake, and groundwater flowing into and out of the lake¹³.

The ice-core record reveals two periods of possible agricultural activity in the period from AD 470 to 1000. The characteristics of the two dust events, however, differ greatly. The AD 490–620 event (Fig. 2) shows a gradual increase in dust which peaks in AD 610–620 and then drops abruptly. As inferred from the Quelccaya accumulation and $\delta^{18}\text{O}$ records, climatic conditions at the onset were dry and warm, whereas warm and wet conditions prevailed from AD 580 to 640, with maximum snow accumulation and dust deposition from AD 610 to 620.

In contrast, the AD 830–960 dust event is quite different. The steady increase in dust deposition peaks in AD 920 and gradually decreases until AD 1060. Also, the accumulation remains high throughout, with maximum snow accumulation coinciding with maximum dust influx from AD 910 to 920.

Low-latitude, high-altitude glaciers offer a tremendous potential for high-temporal-resolution pollen studies, as they are located in close proximity to abundant plant life. Quelccaya ice-core samples typically contain several hundred to over a thousand grains of pollen and spores per litre of water, which is two orders of magnitude higher than that found in the Devon ice cap in the Canadian Arctic¹⁴. In these determinations, water samples from the ice cores were filtered through 7- μm Nitex cloth; the residues were treated with hydrofluoric acid, acetylated, stained and mounted on one slide per sample. One tablet of exotic *Eucalyptus* was added to each water sample before processing for estimating pollen concentration. Except in a few samples, 35–124 fossil grains and 1100–3600 *Eucalyptus* tracers were counted per sample.

Specific pollen samples collected from the AD 920 dust event show no conclusive evidence of agricultural activities. Although pollen concentrations are relatively high (~1,600 grains per litre), the percentages of major pollen taxa (Gramineae, Compositae, *Plantago*, *Alnus*) are not significantly different from those of adjacent samples. Unfortunately, most of the plants cultivated on the altiplano do not produce wind-blown pollen, and thus pollen cannot be used in this case to identify specific crops grown in the raised fields.

The ice cores from the Quelccaya ice cap contain a 1,500-yr record of the climatic variability that affected human activities in southern Peru. The dust peaks from before AD 1000 may reflect agricultural and grazing activity, climatic change, or most probably a combination of the two. Information recovered from ice cores provides a valuable supplement to the archaeological record and may help investigators uncover the patterns and timing of some pre-Hispanic cultural activities in South America. Additional ice-core and archaeological studies must be conducted in the tropical Andes to determine whether or not these two dust events are local and unique to the altiplano of Southern Peru or whether they reflect a larger-scale regional occurrence of an as yet unknown cause.

This work was supported by NSF Division of Atmospheric Sciences and the Division of Polar Programs. We thank Professor W. Dansgaard and Dr P. Grootes for the oxygen isotope analyses of summit core and core 1, and Dr Santiago E. Antunez de Mayolo for water level changes in Lake Titicaca. We thank numerous scientists, engineers and technicians from Electropetu for scientific and logistical assistance and the Inter-American Geodesic Survey office in Lima for logistical support. Many people have participated in the ten field seasons and their efforts are gratefully acknowledged.

Received 12 August; accepted 10 November 1988.

1. Thompson, L. G., Mosley-Thompson, E., Bolzan, J. F. & Koci, B. R. *Science* **229**, 971 (1985).
2. Thompson, L. G., Mosley-Thompson, E., Dansgaard, W. & Grootes, P. M. *Science* **234**, 361 (1986).
3. Thompson, L. G. & Mosley-Thompson, E. in *Abrupt Climatic Change: Evidence and Implications* (eds Berger, W. H. & Labeyrie, L. D.) 99–110 (NATO ASI Series, 1987).
4. Kolata, A. L. *Am. Antiquity* **51**, 748–762 (1986).
5. Bastien, J. W. *Am. Ethnological Soc. Monogr.* **64**, (1978).
6. Cardich, A. in *Advances in World Archeology* (eds Wendorf, F. & Close, A. E.) 193–333 (Academic, New York, 1985).
7. Denevan, W. M. *Science* **169**, 647–654 (1970).
8. Smith, C. T., Denevan, W. M. & Hamilton, P. *Geogr. J.* **134**, 354–367 (1968).
9. Kolata, A. L. *Field Museum nat. Hist. Bull.* **53**, 13–28 (1982).
10. Kolata, A. L. in *Ancient South Americans*, (eds Jennings, J. D. & Freeman, W. H.) 241–284 (San Francisco, 1983).
11. Thompson, L. G., Mosley-Thompson, E. & Arno, B. M. *Science* **226**, 50–53 (1984).
12. Künzel, F. & Kessler, A. *Archs Met. Geophys. Bioclim. Ser. B* **36**, 219–227 (1986).
13. Street-Perrott, F. A. & Harrison, S. P. in *Paleoclimatic Analysis and Modelling* (ed. Hecht, A. D.) 291–340 (Wiley, New York, 1985).
14. McAndrews, J. H. *Quat. Res.* **22**, 68–76 (1984).

Blindsight and insight in visuo-spatial neglect

John C. Marshall & Peter W. Halligan*

Neuropsychology Unit, University Department of Clinical Neurology, Oxford OX2 6HE, UK

* Rivermead Rehabilitation Centre, Oxford OX1 4XD, UK

In a variety of neurological syndromes, patients may show tacit awareness of stimuli that cannot be consciously recollected or identified¹. Such dissociations are the defining characteristic of 'blindsight'^{2,3}; comparable phenomena are seen in some patients with amnesia⁴ and some with prosopagnosia, a profound impairment of familiar face recognition⁵. We report here an analogous dissociation between overt and covert perception in a case of visuo-spatial neglect⁶. The patient, P.S., had sustained right cerebral damage and failed overtly to process information in the hemispace contralateral to lesion. In common with most patients who manifest left-sided neglect, P.S. has a left homonymous hemianopia. Nonetheless, her neglect persists despite free movement of the head and eyes and is thus not a direct consequence of sensory loss in the left visual field. P.S. was presented simultaneously with two line drawings of a house, in one of which the left side was on fire. She judged that the drawings were identical; yet when asked to select which house she would prefer to live in, she reliably chose the house that was not burning.

P.S., a 49-year-old woman, sustained a subarachnoid haemorrhage, confirmed by CT scan and angiography, from an aneurysm at the bifurcation of the basilar artery. Right fronto-temporal craniotomy was performed with clipping of the neck and basilar aneurysm; post-operatively, the patient developed left hemiparesis, and left homonymous hemianopia with macular sparing. On neuropsychological examination, the only finding of note was florid left neglect. Presented in free vision with arrays of simple figures to cancel, P.S. crossed out the stimuli on the right-hand side of the page, neglecting all those on the left. Requested to bisect horizontally-oriented lines (11 inches long) centred on the midsagittal plane, she typically placed her transections over 50% to the right of true centre. When copying simple line drawings and drawing from memory, P.S. made accurate representations of the right side of the object but omitted the left side without any ('conscious') awareness that the drawing was inadequate. On reading individual words, P.S. would frequently omit or substitute the leftmost letters (simile → "mile"; facade → "arcade")⁷. On traditional clinical criteria, then, P.S. manifested prototypical left visuo-spatial neglect⁶.

The following observations were made during one testing session, 21 weeks post-operation. The stimuli were four dark green line drawings of a house, in two of which bright red flames were emerging from the left and right sides of the house, respectively. Each stimulus card measured 138 mm by 130 mm. The experiment proceeded in six steps.

Step 1: the examiner indicated that he had a set of cards. P.S. was then shown individually two (otherwise matched) cards, one of which had flames on the left and one which did not. Upon each presentation, she was asked to describe the drawing. This was repeated twice, and each time, P.S. responded "A house".

Step 2: the two cards (vertically aligned) were placed on the desk top in front of the patient in free vision, and P.S. was asked if the two houses were "the same or different". She replied that they were the same. She was then asked if there was "anything wrong" with either card. She replied "no" and repeated that "they are the same".

Step 3: the cards were again presented together on the desk top, one above the other and centred on the midsagittal plane

of the patient's head and trunk. The viewing distance was approximately 16 inches, although body, head and eye movements were in no way constrained. This was repeated 11 times, with alternate vertical ordering. On each trial, P.S. was asked which house she would prefer to live in. She thought this a silly question ("because they're the same"), but when forced to make a response, chose the non-burning house on 9/11 trials ($P = 0.033$, binomial test).

Step 4: the cards were again presented together, one above the other, for six trials, with vertical positioning once more alternated. After each trial, P.S. was first asked whether the two stimuli were the same or different, and then requested to say which house she would prefer to live in. On the first occasion, she responded that the house with flames was slightly bigger than the other; on the second, she thought the house without flames was bigger. She never noticed the flames, and in the four subsequent trials said that the houses were identical. On 5/6 trials, she preferred the non-burning house ($P = 0.109$). Summing the preferences from steps 3 and 4, the probability associated with the null hypothesis (no reliable preference) is $P = 0.006$.

Step 5: two new cards, one with flames on the right side of the house and one normal house, were presented simultaneously. Their respective vertical positions were alternated over six trials. P.S. was again asked on each trial if the cards were the same or different and then asked which house she preferred. On all six trials, she immediately noticed the flames, said that she preferred the non-burning house ($P = 0.016$), and finally remarked, "I hope there's a point to all this..".

Step 6: we then reverted to the conditions of presentation of step 4 (flames on the left). On the first four trials, P.S. said that the paired stimuli were identical, and on 3/4 occasions again preferred the non-burning house. Summing these preferences from steps 3, 4 and 6 we can reject the null hypothesis at $P = 0.004$. On trial 5, P.S. exclaimed "Oh my God, this one's on fire"! Insight was thus finally achieved after the patient had been primed by step 5.

These results are similar to those obtained in patients with 'blindsight'⁸ and prosopagnosia⁹ in that P.S. showed tacit awareness of left space (as indexed here by rational non-random preferences) despite explicit denial of any basis upon which to make the requisite choice (the stimuli are judged as identical). The results differ in that P.S. can be cued (step 5) into overt identification of the unrecognized or 'neglected' feature. Failure of conscious perceptual awareness is thus not absolute.

Our findings are compatible with those found in a different paradigm by Volpe and co-workers¹⁰. Their patients, with tumours of the right parietal lobe, had full visual fields. They nevertheless showed extinction of the left space stimulus when, with central fixation, two objects were presented simultaneously, one in the left field and one in the right field. Despite failure to report the left-field stimulus, the patients responded at a level significantly above chance when asked to judge whether two stimuli (one in each field) were the same or different. Like P.S., their patients thought the task "silly" as they could only 'see' and identify one stimulus. The patients¹⁰ were not reported to show neglect (extinction to bilateral simultaneous stimulation can be seen in patients without other signs of visual inattention). Karnath and Hartje¹¹, however, have found closely similar results on one subject with manifest left neglect but no visual field defect. Like our patient, this man bisected lines significantly to the right of centre and omitted the left side when copying simple line drawings. With central fixation, two stimuli, presented for 180 ms 3°–5° left and right of fixation, could be reliably judged as the same or different. Yet on trials where the judgement 'different' was (correctly) elicited, the left visual field stimulus could only be identified on about 50% of presentations. In contrast to these studies using bilateral simultaneous stimulation for brief exposures, our experiment is, as far as we are aware, the first demonstration that analogous results can be obtained

with central presentation, free vision and unlimited viewing time in a patient with severe visuo-spatial neglect.

It would therefore seem that the so-called 'neglected' stimulus can exert an influence upon cognitive functioning, albeit at some pre-attentive, pre-conscious level. Exactly how many dissociations between indirect measures of pre-conscious processing can be discovered in the modular brain is a crucial issue for further research. The patients studied by the other groups^{10,11} could make accurate same/different assessments despite 'extinction' of one member of the judged pair. Our patient could report both members of a pair that differed on a critical feature (flames on the left of one of the houses) but nonetheless judged them identical. Yet her preference judgements clearly indicated covert knowledge of that feature. Mandatory mechanisms of input analysis must have far more extensive pre-conscious access to output systems than are dreamt of in most current theories of brain organization¹².

This work was supported by the Chest, Heart, and Stroke Association (P.W.H.) and by the MRC (J.C.M.). We also thank Dr A. Allport and Mr J. Driver for discussions that led to this study.

Received 24 October; accepted 24 November 1988.

- Schacter, D. L., McAndrews, M. P. & Moscovitch, M. in *Thought Without Language* (ed. Weiskrantz, L.) 242-278 (Clarendon, Oxford, 1988).
- Pöppel, E., Held, R. & Frost, D. *Nature* **243**, 295-296 (1973).
- Weiskrantz, L., Warrington, E. K., Sanders, M. D. & Marshall, J. *Brain* **97**, 709-728 (1974).
- Milner, B., Corkin, S. & Teuber, H.-L. *Neuropsychologia* **6**, 215-234 (1968).
- Bauer, R. M. *Neuropsychologia* **22**, 457-469 (1984).
- Jeannerod, M. (ed.) *Neuropsychological and Neurophysiological Aspects of Spatial Neglect* (North-Holland, Amsterdam, 1987).
- Kinsbourne, M. & Warrington, E. K. *J. Neurol. Neurosurg. Psychiatr.* **25**, 339-344 (1962).
- Weiskrantz, L. *Brain* **110**, 77-92 (1987).
- de Haan, E. H. F., Young, A. W. & Newcombe, F. *Cognitive Neuropsychol.* **4**, 385-415 (1987).
- Volpe, B. T., Ledoux, J. E. & Gazzaniga, M. S. *Nature* **282**, 722-724 (1979).
- Karnath, H.-O. & Hartje, W. *J. Neurol.* **234**, 180-184 (1987).
- Fodor, J. A. *The Modularity of Mind* (MIT, Cambridge, Massachusetts, 1983).

Genetic influence on general mental ability increases between infancy and middle childhood

D. W. Fulker*, J. C. DeFries* & Robert Plomin†

* Institute for Behavioural Genetics, University of Colorado, Boulder, Colorado 80309-0447, USA

† College of Human Development, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Adoption studies can provide direct evidence for the independent effects of family environment and heredity that are always confounded in intact nuclear families¹⁻³. When children are separated from their biological mothers shortly after birth and placed non-selectively in adoptive homes, adoptive-parent/adopted-child resemblance can be ascribed to cultural transmission, whereas biological-parent/adopted-child similarities are due to heritable factors. Furthermore, a longitudinal adoption study facilitates examination of changes in these two main sources of variation during development. The Colorado Adoption Project is the first large-scale longitudinal adoption study of behavioural development and was initiated in 1975⁴. Data were collected from biological parents of 245 adopted children, the adoptive parents and parents of 245 matched nonadopted children. The children have subsequently been tested at 1, 2, 3 and 4 years of age, and at the end of their first year in primary school (average age, 7.4 years). The number of subjects tested is now adequate for analysis of data over 7 years. The results provide conclusive evidence for increasing heritable variation of general mental ability, ranging from 9% at 1 year of age to 36% at 7 years.

The only previous longitudinal adoption study of cognitive development was reported by Skodak and Skeels⁵. IQ (intel-

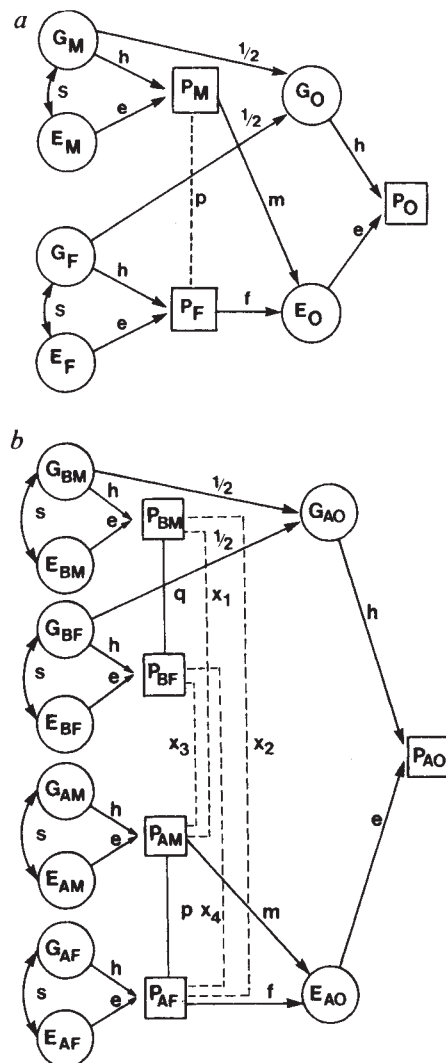


Fig. 1 Path model of cultural and genetic transmission in *a*, nonadoptive and *b*, adoptive families. The relative impact of the latent variables *G* and *E* on the manifest variable *P* is indicated by the path coefficients *h* and *e*. The latent and manifest variables are subscripted M, F and O for mothers, fathers and offspring in nonadoptive families, and BM, BF, AM, AF and AO for biological mothers, biological fathers, adoptive mothers, adoptive fathers and adopted offspring in adoptive families. As a result of the combined effects of cultural and genetic transmission in nonadoptive families, *G* and *E* may be correlated *s*. Phenotypic assortative mating, which induces correlations between the latent variables of the parents, is represented by a conditional path *p* for adoptive and nonadoptive parents, and *q* for biological parents of adopted children¹¹. Genetic transmission is represented by a path of 1/2 from the genotype of each parent, whereas the effects of parental mental ability on the environment of their child are indicated by the paths *m* and *f*.

ligence quotient) scores of 100 children adopted before 6 months of age and tested at least four times between infancy and adolescence were compared with the educational level and occupational status of both their adoptive and biological parents, and with the IQ scores of 63 of the biological mothers. In general, correlations between the IQ scores of biological mothers and their adopted-away children suggested an increasing genetic influence from early childhood to adolescence, but the number of biological mothers was insufficient to assess heritable variation adequately, and there was substantial selective placement (the matching of biological and adoptive parents) for educational level. In the Colorado Adoption Project,

Table 1 Observed parent-child correlations at five different ages

Parental type	Age of child (years)				
	1	2	3	4	7
Biological mother	0.10 (232)	0.07 (211)	0.14 (203)	0.18 (194)	0.31 (156)
Biological father	0.33 (47)	0.35 (44)	0.21 (44)	0.45 (42)	0.12 (35)
Biological parents (pooled)	0.14	0.12	0.15	0.23	0.28
Adoptive mother	0.11 (232)	0.07 (211)	0.15 (203)	0.16 (194)	-0.05 (156)
Adoptive father	0.08 (232)	0.04 (211)	0.20 (203)	0.12 (194)	0.16 (156)
Adoptive parents (pooled)	0.09	0.05	0.16	0.14	0.06
Nonadoptive mother	0.01 (241)	0.14 (227)	0.13 (210)	0.17 (210)	0.16 (138)
Nonadoptive father	0.10 (241)	0.13 (227)	0.12 (210)	0.12 (210)	0.17 (138)
Nonadoptive parents (pooled)	0.05	0.14	0.13	0.14	0.17

Parent-parent correlations for year 1 (when N is largest) are as follows: adoptive mother/adoptive father, 0.29 (232); biological mother/biological father, 0.30 (47); nonadoptive mother/nonadoptive father, 0.25 (241); biological mother/adoptive mother, -0.04 (232); biological mother/adoptive father, 0.04 (232); biological father/adoptive mother, -0.06 (47); biological father/adoptive father, 0.05 (47); pooled assortment correlation, 0.27 (519); pooled selective placement correlation, 0.00. Values of N are given in parentheses.

however, biological mothers and adoptive parents of adopted children and parents of nonadopted children were administered a three-hour battery of psychometric tests, about 20% of the biological fathers were also tested, the sample size was large, and selective placement was minimal.

The sample is representative of metropolitan Caucasian families in the United States with regard to both means and variances of socioeconomic measures⁴. Because biological parents are considerably younger than adoptive or nonadoptive parents, their average socioeconomic status is lower, but the parents of the biological, adoptive and nonadoptive parents are very similar with regard to demographic variables, suggesting that the biological parents of adopted children are not from a disadvantaged group.

A battery of 13 tests of specific cognitive abilities was administered to adults. The tests, their reliability, the methods applied for gender and age adjustment, and factor structure are described elsewhere⁶. The first unrotated principal component score is used as a measure of general mental ability. With regard to the children, the Bayley Scales of Infant Development⁷ were administered at 1 and 2 years of age. The Stanford-Binet Intelligence Scale-Form L-M (ref. 8) was administered at 3 and 4 years, and the Wechsler Intelligence Scale for Children-Revised⁹ was administered during middle childhood (average age, 7.4 years).

In this study we analyse separately the family data for children tested at each age. The number of adoptive families in each analysis ranged from 232 to 156, and the number of nonadoptive families ranged from 241 to 138.

The basic Colorado Adoption Project correlations are shown in Table 1. Certain features of these correlations are apparent. Selective placement correlations shown in the footnote are close to zero, thus permitting interpretation of the biological and adoptive parent-child correlations in terms of the independent effects of genotype and home environment. The biological correlations appear to increase with age, particularly in the maternal data where the N is large, suggesting an increase in heritability. The paternal genetic effect is more variable, as we might expect when N is much smaller. Adoptive parent-child correlations, which reflect the direct effect of parental environmental influence on the child, appear to increase with age until 3 or 4 years, but to decrease at 7 years. The combined effects of genetic and environmental influences should be evident in the nonadoptive parent-offspring correlations. These show little trend, as we would expect, because one component increases and the other rises and then falls, but they do seem anomalously low. When the correlations are evaluated as a whole in the context of the model we develop later, the χ^2 goodness-of-fit shows no statistically significant anomaly however.

Path analysis was used to assess the genetic and environmental

Table 2 Expected correlations among manifest variables in nonadoptive and adoptive families

Variable	Expected correlation
Nonadoptive families	
P_M, P_F	$[p]$
P_M, P_O	$[1/2h(1+p)(h+se) + e(m+fp)]$
P_F, P_O	$[1/2h(1+p)(h+se) + e(f+mp)]$
Adoptive families	
P_{BM}, P_{BF}	$[q]$
P_{BM}, P_{AM}	$[(x_1 + qx_3) + (x_2 + qx_4)p]$
P_{BM}, P_{AF}	$[(x_1 + qx_3)p + (x_2 + qx_4)]$
P_{BM}, P_{AO}	$\sigma_{P_O}/\sigma_{P_{AO}}[em\{(x_1 + qx_3) + (x_2 + qx_4)p\} + ef\{(x_1 + qx_3)p + (x_2 + qx_4)\} + 1/2h(1+q)(h+se)]$
P_{BF}, P_{AM}	$[(qx_1 + x_3) + (qx_2 + x_4)p]$
P_{BF}, P_{AF}	$[(qx_1 + x_3)p + (qx_2 + x_4)]$
P_{BF}, P_{AO}	$\sigma_{P_O}/\sigma_{P_{AO}}[em\{(qx_1 + x_3) + (qx_2 + x_4)p\} + ef\{(qx_1 + x_3)p + (qx_2 + x_4)\} + 1/2h(1+q)(h+se)]$
P_{AM}, P_{AF}	$[p]$
P_{AM}, P_{AO}	$\sigma_{P_O}/\sigma_{P_{AO}}[e(m+fp) + 1/2h(h+se)\{(x_1 + qx_3) + (x_2 + qx_4)p + (qx_1 + x_3) + (qx_2 + x_4)p\}]$
P_{AF}, P_{AO}	$\sigma_{P_O}/\sigma_{P_{AO}}[e(f+mp) + 1/2h(h+se)\{(x_1 + qx_3)p + (x_2 + qx_4) + (qx_1 + x_3)p + (qx_2 + x_4)\}]$

There are three constraints among the parameters in the model. The standardized phenotypic variance for nonadoptive parents and their children is $h^2 + 2hse + e^2 = 1.0$, whereas that for adopted children is $h^2 + he\{(x_1 + qx_3) + (x_2 + qx_4)p + (qx_1 + x_3) + (qx_2 + x_4)p\} + \{(x_1 + qx_3)p + (x_2 + qx_4) + (qx_1 + x_3)p + (qx_2 + x_4)\}f + e^2$, which reduces to $h^2 + e^2$ in the absence of selective placement. Moreover, $s = 1/2\{(1+p)(h+se)(m+f)\}$. Subscript notation and parameters in Fig. 1, see legend.

aetiologies of individual differences in general mental ability during development from infancy to middle childhood. Observed covariance matrices were computed and then equated to their expectations using a maximum-likelihood estimation procedure¹⁰. Three covariance matrices were available at each age: one 5×5 matrix involving all available biological and adoptive parents and the adopted child; one 4×4 matrix in which data on biological fathers are missing; and one 3×3 matrix for nonadoptive mother, father and child. Thus, a total of 31 statistics was available for evaluating the model at each age. A simple linear model was assumed in which latent variables representing additive genetic values (G) and environmental deviations (E) cause variation in phenotypic values (P). This basic model is depicted as a path diagram in Fig. 1a for nonadoptive families, in which parents transmit both hereditary and environmental influences to their children. A simplification in this model was forced on us by the design of the study, and that is the assumption that the path coefficient h is the same for parent and child. If h^2 changes with age, this assumption is false. But if we interpret the genotypes in Fig. 1 as representing genetic factors or influences common to parent and child, then changes in estimates of h^2 in our model indicate changes in the importance of this factor with age.

As shown in Fig. 1b, biological and cultural transmission are dissociated in adoptive families. In order to assess the importance of selective placement and its potential effects on parent-child resemblance, the matching between biological and adoptive parents was modelled by the conditional paths x_{1-4} (ref. 11). Expected correlations among the manifest variables are presented in Table 2. There are 11 parameters in these expectations, two of which (s and e) are functions of the others, as shown. In addition, σ_p^2 for children and σ_p^2 for adults are required to transform the expected correlations into covariances. When selective placement is negligible, as it is here, the x_{1-4} terms may be omitted. In this case the expectations take the simple form characteristic of an ideal adoption study, in which the genotype and environment of adopted children are independent.

The model was fitted to data from parents and children at each of the five ages. Initially, the full model was fitted to estimate all parameters. The χ^2 goodness-of-fit had 31 observations minus 11 parameters, or 20 degrees of freedom for the full model. A series of reduced models was then fitted to assess the significance of the omitted parameters. The four selective placement parameters were first dropped from the model. At 1 year of age, the change in χ^2 from the full to the reduced model was nonsignificant ($\chi^2=0.35$, d.f.=4, $P>0.50$), indicating the absence of selective placement. As expected, corresponding tests for selective placement at subsequent years of age were also nonsignificant. Assortative mating parameters for wedded and unwedded parents ($p=0.27$ and $q=0.28$) were then equated, again with no loss of fit. Similarly, maternal and paternal cultural transmission parameters (m and f) were equated with no significant loss of fit (largest χ^2 change = 2.68, d.f.=1, $P>0.10$ at 7 years).

Parameter estimates resulting from the fit of this parsimonious model to the data are presented in Table 3. With the number of parameters reduced from 11 to 5 for this reduced model, the χ^2 goodness-of-fit statistic has 26 degrees of freedom. The fits of the model at the various ages are adequate, and the parameter estimates manifest an interesting developmental pattern: estimates of h^2 increase from 0.09 to 0.36. To assess the significance of the differences among these estimates of h^2 , we fitted a model of linear trend using the inverse of their sampling variances as weights in a least-squares analysis. Because the errors of these h^2 estimates are positively correlated, this procedure is quite conservative¹². The $\chi^2=0.939$, d.f.=3, $P>0.5$ for linear trend indicates a very good fit. Dropping the slope from the model gave a change in $\chi^2=6.16$, d.f.=1, $P\approx 0.02$. Thus, this conserva-

Table 3 Parameter estimates resulting from the fit of a parsimonious model of genetic and cultural transmission to general cognitive ability data in the Colorado Adoption Project

Parameter	Age in years				
	1	2	3	4	7
h^2	0.09	0.14	0.10	0.20	0.36
$m=f$	0.04	0.05	0.10	0.07	0.01
$p=q$	0.27	0.25	0.27	0.27	0.21
s	0.02	0.03	0.05	0.04	0.01
χ^2	33.52	32.86	29.84	30.13	25.48
d.f.	26	26	26	26	26
$P>$	0.10	0.10	0.20	0.20	0.40

tive test clearly indicates an increase in h^2 with age, as do the raw correlations in Table 1.

The heritability estimates in Table 3 are remarkably congruent with those obtained by doubling the difference between identical and fraternal twin correlations reported for the Louisville Twin Study¹³, namely 0.10, 0.16, 0.18, 0.24 and 0.34 at 1, 2, 3, 4 and 8 years of age respectively. The similarity of these estimates suggests that genetic effects on mental ability during infancy and childhood are highly correlated with those expressed during adulthood. When Eaves *et al.*¹⁴ fitted a longitudinal genetic model to the Louisville Twin Study data, they also found evidence for increasing heritability with age and a very high genetic correlation from age to age.

In contrast to the pattern of increasing heritable influence, the paths representing parental cultural transmission increase from 0.04 to 0.10 between 1 and 3 years of age, but then decrease to 0.01 at 7.4 years. As expected, given the limited evidence for cultural transmission, estimates of genotype-environment correlation (s) are also small. To test for the significance of cultural transmission, $m=f$ was dropped from the reduced model with a significant loss of fit only at 3 years (change in $\chi^2=7.16$, d.f.=1, $P<0.01$). In the same way, the significance of heritable variation in the presence of cultural transmission was assessed by retaining $m=f$ in the model and dropping h^2 . Resulting changes in χ^2 indicated significant losses of fit at 2, 4 and 7 years (corresponding changes in χ^2 were 3.91, 6.84 and 15.78, with d.f.=1, and $P<0.05$, 0.01 and 0.001 respectively).

Results of the Colorado Adoption Project provide the best evidence to date for the increasing importance of the genetic aetiology of individual differences in mental ability between infancy and middle childhood. In contrast, the direct effects of parental intelligence on the child's environment, especially after one year in primary school, appear to be relatively small.

This research is supported in part by grants from the National Institute of Child Health and Human Development, the National Science Foundation, and the Spencer Foundation. We thank Rebecca G. Miles for editorial assistance.

Received 10 May; accepted 7 November 1988.

1. Cavalli-Sforza, L. L. in *Developmental Human Behavior Genetics* (eds Schaie, K. W., Anderson, V. E., McClearn, G. E. & Money, J.) (Lexington, 1975).
2. Kamin, L. In *The Intelligence Controversy* (eds Eysenck, H. J. & Kamin, L.) (Wiley-Interscience, New York, 1981).
3. Lewontin, R. C. *A. Rev. Genet.* **9**, 387-405 (1975).
4. Plomin, R. & DeFries, J. C. in *Origins of Individual Differences in Infancy: The Colorado Adoption Project* (Academic, Orlando, 1985).
5. Skodak, M. & Skeels, H. M. *J. Genet. Psychol.* **75**, 85-125 (1949).
6. DeFries, J. C., Plomin, R., Vandenberg, S. G. & Kuse, A. R. *Intelligence* **5**, 245-277 (1981).
7. Bayley, N. in *Manual for the Bayley Scales of Infant Development* (Psychological Corporation, New York, 1969).
8. Terman, L. M. & Merrill, M. A. in *Stanford-Binet Intelligence Scale: 1972 Norms Edition* (Houghton Mifflin, Boston, 1973).
9. Wechsler, D. L. in *Examiner's Manual: Wechsler Intelligence Scale for Children-Revised* (Psychological Corporation, New York, 1974).
10. Fulker, D. W. & DeFries, J. C. *Br. J. math. & stat. Psychol.* **36**, 175-188 (1983).
11. Carey, G. *Am. J. hum. Genet.* **39**, 775-786 (1986).
12. McGue, M., Wette, R. & Rao, D. C. *Genet. Epidemiol.* **1**, 255-269 (1984).
13. Wilson, R. C. *Child Dev.* **54**, 298-316 (1983).
14. Eaves, L. J., Long, J. & Heath, A. C. *Behav. Genet.* **16**, 143-162 (1986).

The role of sex steroids in the acquisition and production of birdsong

Peter Marler, Susan Peters, Gregory F. Ball,
Alfred M. Duffy Jr & John C. Wingfield

The Rockefeller University Field Research Center, Tyrrel Road,
Millbrook, New York 12545, USA

Male birdsong is generally regarded as a secondary sexual characteristic under the control of gonadal steroids^{1,2}. Song typically waxes and wanes with the seasonal cycle of testicular growth and regression and decreases after adult castration. Testosterone therapy reinstates song, induces it in females, augments it in intact males, and spring testosterone profiles correlate with seasonal song production³⁻¹⁰. Thus, testosterone has been viewed as a major factor in song acquisition and production^{3,11-14} acting either directly, or after aromatization within the brain¹⁵⁻¹⁶. We show here, however, that song learning and early phases of the development of singing both take place in castrated male birds with no significant levels of testosterone in their blood plasma. Testosterone seems to be required for song crystallization, however. Oestradiol was unexpectedly still present after castration, evidently from a non-testicular source, throughout the period of male song acquisition.

Male songbirds were reared by hand, castrated at three weeks (10 swamp sparrows, *Melospiza georgiana*) or four weeks (seven song sparrows, *M. melodia*) of age, housed separately in sound-insulated chambers, kept on natural photoperiods, and tape-recorded weekly. Each day during their first year they heard a programme of tape-recorded songs, which changed every 1-6 weeks, totalling 60 song types in all. Blood samples were taken monthly in midwinter, and otherwise every two weeks. Testosterone and 17 β oestradiol levels were measured by radioimmunoassay, after partial purification of the steroid fractions by

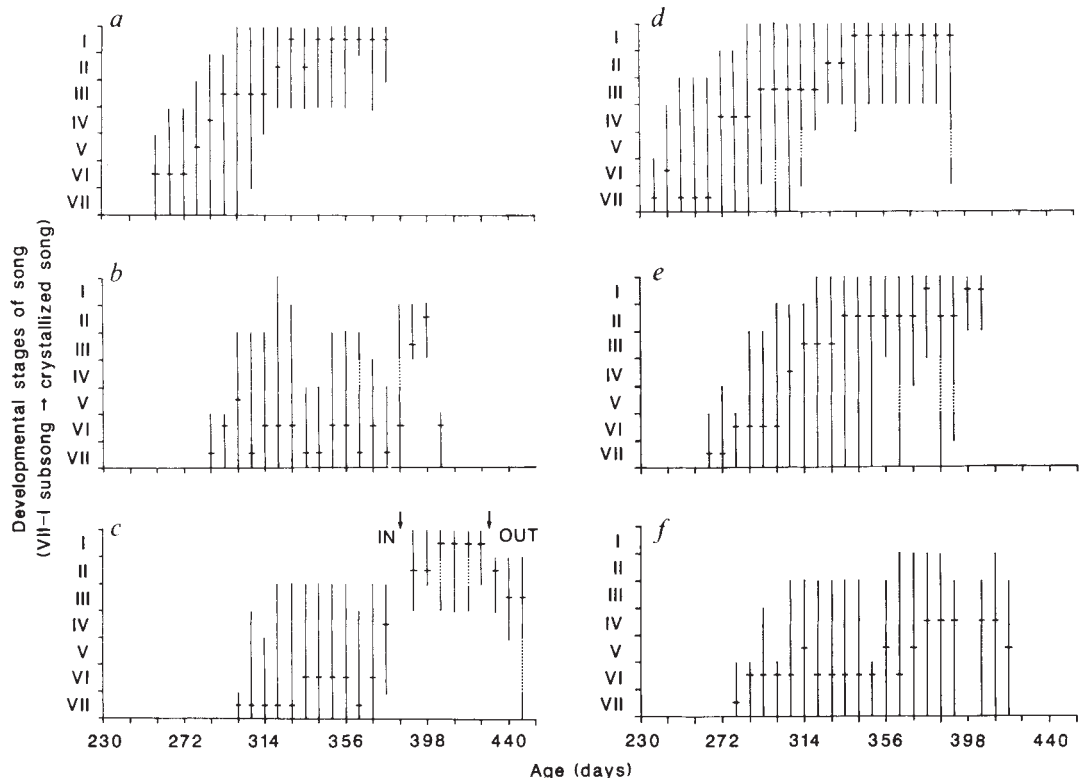
column chromatography^{7,17}. For both hormones this assay, which has been reliably used with sparrows for 15 years, has been fully validated for parallelism, precision, accuracy and sensitivity¹⁷. All songs were analysed by sound spectrography, classified into seven developmental stages from subsong to crystallized song, and examined for evidence of imitation of song models presented at various ages¹⁸. For comparison, six intact male swamp sparrows were reared and trained with tape-recorded song under similar conditions. The song sparrow controls were seven intact males reared from the egg, and treated similarly, but exposed to tape-recorded songs for 10 weeks, beginning when they were 9-17 days old. The castrates were checked for residual testicular tissue, first by laparotomy, and then during autopsy when two years old. All castrations seemed complete except for three song sparrows, called incomplete castrates.

All seven castrated song sparrows began song development about one month later than controls. Song development proceeded in a similar manner in castrates and controls (Fig. 1). Rates of singing during the plastic song stage were comparable, but only two birds gave crystallized song, both incomplete castrates. One produced normal crystallized song, and another progressed from subsong to the most advanced plastic song phase, and gave a few crystallized songs. Song development in the incomplete castrates was intermediate between that for controls and the complete castrates in timing and stages achieved (Fig. 1d-f). The complete castrates sang with decreasing frequency, and regressed to earlier stages after about 400 days (Fig. 1). Analysis revealed that five castrates (two complete, three incomplete) imitated tape-recordings (Fig. 2), mostly (80%) acquired between 30 and 100 days of age, as in intact song sparrows¹⁹. In the first year two males failed to progress to plastic song, when imitations are first detectable.

The 10 swamp sparrow castrates also began singing late. Early stages of song development were normal, but subsong tended to persist in the castrates, instead of ceasing as more advanced song emerged. Whereas crystallization occurred at about 330

Fig. 1 Patterns of song development in castrated and intact 230-447-day-old male song and swamp sparrows. *a*, Intact swamp sparrows ($N=6$). *b*, Castrated swamp sparrows ($N=5$). *c*, Castrated swamp sparrows + testosterone ($N=5$). *d*, Intact song sparrows ($N=10$). *e*, Song sparrows: incomplete castrates ($N=3$). *f*, Song sparrows: complete castrates ($N=4$).

Methods. Sound spectrograms of weekly taped-recordings were analysed into seven stages, from subsong (VII, VI), a soft, amorphous warbling, through plastic song (V-II), in which mature song structure first appears in variable form, with imitations, to crystallized song (I), with the loud, repetitive, stereotyped patterns of mature song. The bar for a given age indicates the range from the least advanced stage achieved, and the median stage averaged for all subjects. Treatment of the 10 castrated male swamp sparrows was identical until 384 days, when five were implanted with a silastic capsule of testosterone for six weeks.



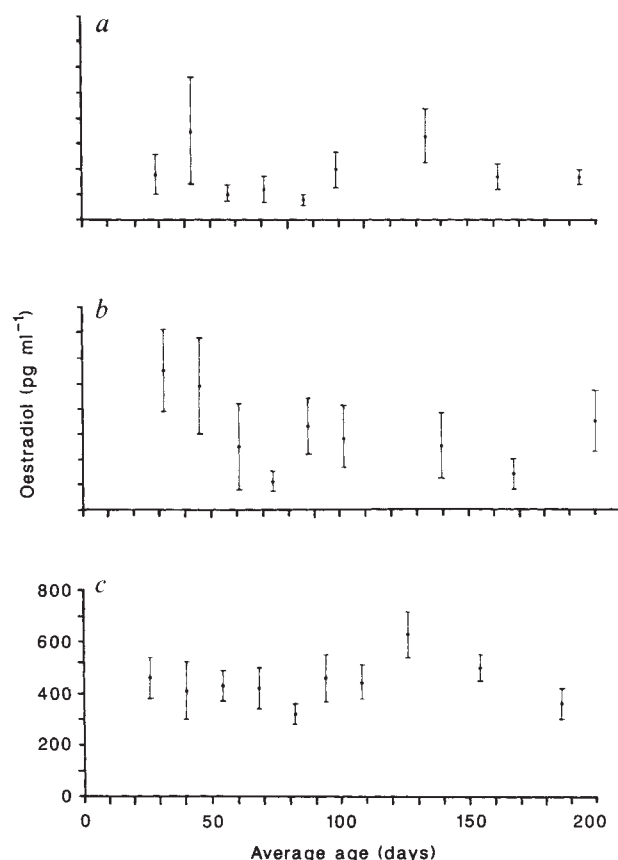


Fig. 2 Plasma levels of oestradiol (means and standard errors) in intact male song sparrows (top) and castrated male song (middle) and swamp (bottom) sparrows. Data are presented from 20–200 days of age, after which oestradiol ceased to be detectable.

days of age in intact birds¹⁸, only one castrate produced a single bout of crystallized song at 324 days, and then sang sporadic plastic song up to about 400 days of age. The other nine birds produced only subsong and some early plastic song (Fig. 1b).

At about 380 days of age, the swamp sparrows were divided into five matched pairs, on the basis of stages of vocal development then achieved. One member of each pair was implanted with testosterone for six weeks. Within 1–3 weeks, the five testosterone-treated birds produced crystallized song, and continued until the end of testosterone therapy (Fig. 1c). The five untreated castrates persisted with subsong, or ceased singing altogether. When the testosterone-induced crystallized songs were analysed (five males, 10 song types), six were clear imitations of training songs by three males (Fig. 2), mostly (80%) acquired at the normal time between about 25 and 60 days of age²⁰. At two years of age all castrates of both species were brought into crystallized song again with testosterone implants. Song analysis showed that, of the castrates, all song sparrows and six swamp sparrows had learned in the first year.

Two previous studies have demonstrated vocal acquisition and development in castrated male songbirds^{21,22}. Again, birds failed to crystallize songs normally, but hormone levels were not checked. It has been assumed that the castrations were incomplete or that non-testicular testosterone was responsible. We were able to confirm conclusively that most gonadectomies in our study were successful. In both species, plasma levels of testosterone in all complete castrates remained at or below the level of detectability (200 pg ml⁻¹) throughout the study, as compared with levels in early spring of 1,500 pg ml⁻¹ or more in controls. Testosterone was present in the plasma of the three incompletely castrated song sparrows, at levels below those for

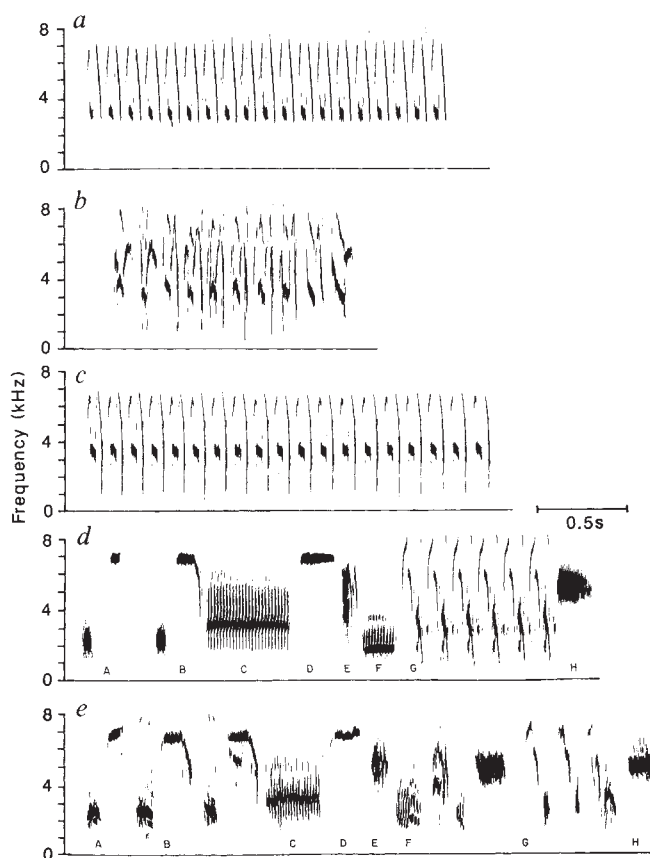


Fig. 3 Sound spectrograms of songs produced by castrated male sparrows. A normal tape-recorded swamp song (a) was played to a castrated male swamp sparrow from 34 to 75 days. b, The most advanced plastic song produced by this male at 376 days, illustrating its variable structure, with hints of imitations. c, A crystallized imitation of a produced by the same castrated male after testosterone therapy. A castrated male song sparrow presented with a normal song sparrow song (d) from 43–85 days produced a plastic song without androgen therapy at 363 days (e) incorporating imitations of the eight components identified (A–H).

controls. High levels were recorded only in the five castrates given androgen therapy, bringing them into the natural physiological range of plasma levels for this season.

Unexpectedly, elevated levels of oestradiol were found in the castrates of both species up to six months of age (Fig. 2b, c), similar to those found in both sets of controls. Values were close in complete and incomplete castrates, so the data were combined (Fig. 2b). Oestradiol has been found in the plasma of male birds^{3,11,23}, although the identification should be confirmed by further chromatographic methods, and by mass spectrometry. In the present case it will be important to verify that the hormone present in castrates is indeed oestradiol, identical with that in the plasma of intact males. Until now, attention has focused on pre- and perinatal effects of oestradiol in preparing the male brain for song production^{24–27}. This is the first indication of possible involvement in the actual process of song learning in later life. The period of oestradiol elevation overlaps completely with the times of song acquisition. In a previous study of song learning in intact male swamp sparrows, we found that oestradiol levels correlated significantly better with readiness to learn than testosterone levels⁷. Thus oestradiol may be implicated in activation of mechanisms underlying song acquisition during adolescence. The general similarity of oestradiol plasma profiles in castrated and intact birds, and in the complete and incomplete castrates, suggests that the normal source of this hormone in male sparrows is non-testicular.

These experiments are the first to dissociate the roles of sex steroids in different phases of the song-learning process. They show definitively that high plasma levels of testosterone are not required, either for song acquisition, or for early stages of song production. Testosterone seems to trigger song crystallization, and is probably necessary to achieve full song. Both the syrinx and song control brain nuclei have receptors for testosterone which, directly or indirectly, affects cell size and number, induces protein synthesis, and encourages dendritic growth and synaptogenesis in several of the nuclei^{2,25,28-30}. The work presented here suggests that these testosterone-related neural events are more closely associated with motor aspects of singing behaviour than with the process of memorizing learned songs.

We are indebted to Judith Marler and Eileen McCue for raising the birds and Drs Arthur Arnold and Stephen Nowicki for criticism of the manuscript. Esther Arruza prepared the figures and manuscript. We thank the Institute of Ecosystem Studies, the New York Botanical Garden at the Mary Flagler Cary Arboretum for access to study sites. Research was supported in part by the Biomedical Research Support Grant Program, Division of Research Resources, NIH, and by an NIMH grant to P.M.

Received 12 September; accepted 8 November 1988.

1. Andrew, R. J. in *Bird Vocalizations* (ed. Hinde, R. A.) 97-130 (Cambridge University Press, Cambridge, 1969).
2. Nottebohm, F. in *Handbook of Physiology—The Nervous System V* (ed. Plum, F.) 85-108 (American Psychological Association, Washington, 1987).

3. Pröve, E. in *Hormones and Behaviour in Higher Vertebrates* (eds Balthazart, J., Pröve, E. & Gilles, R.) 368-374 (Springer, Berlin, 1983).
4. Arnold, A. P. *J. exp. Zool.* **191**, 309-326 (1975).
5. Balthazart, J. in *Avian Biology*, Vol. 7 (eds Farner, D. S., King, J. R. & Parkes, K. C.) 221-365 (Academic, New York, 1983).
6. Heid, P., Güttinger, H. R. & Pröve, E. *Z. Tierpsychol.* **69**, 224-236 (1985).
7. Marler, P., Peters, S. & Wingfield, J. *J. Neurobiol.* **18**, 531-548 (1987).
8. Pohl-Apel, G. & Sossinka, R. *Z. Tierpsychol.* **64**, 330-336 (1984).
9. Silver, R., O'Connell, M. & Saad, R. in *Endocrine Control of Sexual Behavior* (ed. Beyer, C.) 223-278 (Raven, New York, 1979).
10. Wingfield, J. C. in *Avian Endocrinology—Environmental and Ecological Perspectives* (eds Mikami, S. I., Homma, K. & Wada, M.) 265-288 (Japan Scientific, Tokyo, and Springer, Berlin, 1983).
11. Nottebohm, F., Nottebohm, M. E., Crane, L. A. & Wingfield, J. C. *Behav. Neural Biol.* **47**, 197-211 (1987).
12. DeVogd, T. J. *J. Neurobiol.* **17**, 177-201 (1986).
13. Nottebohm, F. *Ibis* **111**, 386-387 (1969).
14. Nottebohm, F. in *Progress in Psychobiology and Physical Psychology*, Vol. 9, (eds Sprague, J. M. S. & Epstein, A. N. E.) 85-124 (Academic, New York, 1980).
15. Harding, C. in *Hormones and Behaviour in Higher Vertebrates* (eds Balthazart, J., Pröve, E. & Gilles, R.) 275-289 (Springer, Berlin, 1983).
16. Hutchison, J. B. & Steimer, Th. in *Hormones and Behaviour in Higher Vertebrates* (eds Balthazart, J., Pröve, E. & Gilles, R.) 261-273 (Springer, Berlin, 1983).
17. Wingfield, J. C. & Farner, D. S. *Steroids* **26**, 311-327 (1975).
18. Marler, P. & Peters, S. *Auk* **99**, 446-458 (1982).
19. Marler, P. & Peters, S. *Ethology* **76**, 89-100 (1987).
20. Marler, P. & Peters, S. *Ethology* **77**, 76-84 (1988).
21. Arnold, A. P. *J. exp. Zool.* **191**, 261-278 (1975).
22. Kroodsmä, D. E. *Anim. Behav.* **34**, 1572-1575 (1986).
23. Weichel, K., Schwager, G., Heid, P., Güttinger, H. R. & Pesch, A. *Ethology* **73**, 281-294 (1986).
24. Adkins-Regan, E. in *Hormones and Behaviour in Higher Vertebrates* (eds Balthazart, J., Pröve, E. & Gilles, R.) 218-228 (Springer, Berlin, 1983).
25. Arnold, A. P., Bottjer, S. W., Brenowitz, E. A., Nordeen, E. J. & Nordeen, K. W. *Brain Behav. Evol.* **28**, 22-31 (1986).
26. Hutchison, J. B., Wingfield, J. C. & Hutchison, R. E. *J. Endocr.* **103**, 363-369 (1987).
27. Konishi, M. & Gurney, M. E. *Trends Neurosci.* **5**, 20-23 (1982).
28. Nottebohm, F. *Brain Res.* **189**, 429-436 (1980).
29. Konishi, M. & Akutagawa, E. *Brain Res.* **222**, 442-446 (1981).
30. Nottebohm, F. *Science* **214**, 1368-1370 (1981).

A community effect in animal development

J. B. Gurdon

Cancer Campaign Molecular Embryology Unit, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

In animal development, the first tissues to be formed include such major components as muscle, nerve cord, notochord and the eye. In the vertebrates, all of these tissues are formed by embryonic induction, a process by which some of the cells within a mass of tissue are caused to change their direction of differentiation as a result of close proximity to cells of another kind¹⁻³. The induced cells typically form a solid coherent mass with a distinct border between them and the remaining uninduced cells (Fig. 1a). This clean separation between induced and uninduced cells is much sharper than can readily be explained as a result of the induction process. We describe here the culture of amphibian cell and tissue recombinations in solid gels containing cytochalasin in which cell division and cell movement is inhibited during response to induction. This has revealed an effect in which the ability of a cell to respond to induction by differentiating as muscle is enhanced by, or even dependent on, other neighbouring cells differentiating in the same way at the same time. This seems to be a newly described process in animal development, termed the community effect. It helps to explain the formation of blocks of tissue from sheets of cells, and could be of widespread occurrence and significance in morphogenesis resulting from embryonic induction.

The experimental system used here is that in which vegetal cells of a *Xenopus laevis* blastula induce cells of the animal hemisphere to differentiate as muscle⁴⁻⁶; in the absence of induction, these animal cells would have formed epidermis, expressing cytokeratin genes⁷⁻⁹. An essential feature of this induction system is that the induced muscle cells are typically concentrated together in a solid block of tissue sharply demarcated from adjacent non-muscle cells (Fig. 1b). This

experimental system therefore reproduces the normal morphogenetic process being studied.

Evidence for a community effect emerged from attempts to induce a single animal cell placed on vegetal tissue to undergo muscle differentiation. Many efforts of this kind were unsuccessful, but this could be explained as an adverse effect of the culture medium on a single exposed animal cell; when blocks of animal tissue containing several hundred cells are induced by vegetal cells, it is the internal animal cells, not exposed to the medium, which differentiate as muscle. In view of this, a sandwich construct has been designed to analyse the response of single animal cells to induction. Animal regions of a blastula have been fully dissociated, in Ca²⁺-free medium, into single cells. These are then reaggregated, by the addition of calcium ions, into either monolayer or solid reaggregates, and the reaggregated cells introduced in one or other of these configurations into a sandwich composed of the vegetal pieces of two blastulae (Fig. 2). To permit the subsequent distinction between animal and vegetal cells, the animal cells, but not vegetal cells, were prelabelled with [³H]thymidine or [³H]succinimidyl-propionate¹⁰. The sandwich conjugates were cultured in a cytochalasin-containing gel (the gel prevents the cells falling apart) for another day until muscle gene expression can be detected by antibodies¹¹. At this time, conjugates were fixed, serially sectioned, reacted with a muscle-specific antibody and Hoechst nuclear stain, and autoradiographed¹². Each cell of a conjugate was then scored as being of animal or vegetal origin (autoradiographic labelling), of normal nuclear structure (Hoechst), and as expressing muscle-specific genes or not (fluorescent antibody). In this experimental design, all animals cells are protected from the culture medium, and all are treated in the same way except for being reaggregated into a monolayer or solid configuration.

The results of these analyses, illustrated in Fig. 3, are that animal cells in monolayer configuration do not activate muscle genes, whereas a high proportion of the same cells as solid reaggregates do. Table 1 summarizes the results of many such experiments, and shows that, if cells with condensed nuclei are excluded, a large majority of solid reaggregate cells differentiate as muscle, whereas none of the same cells in a monolayer do

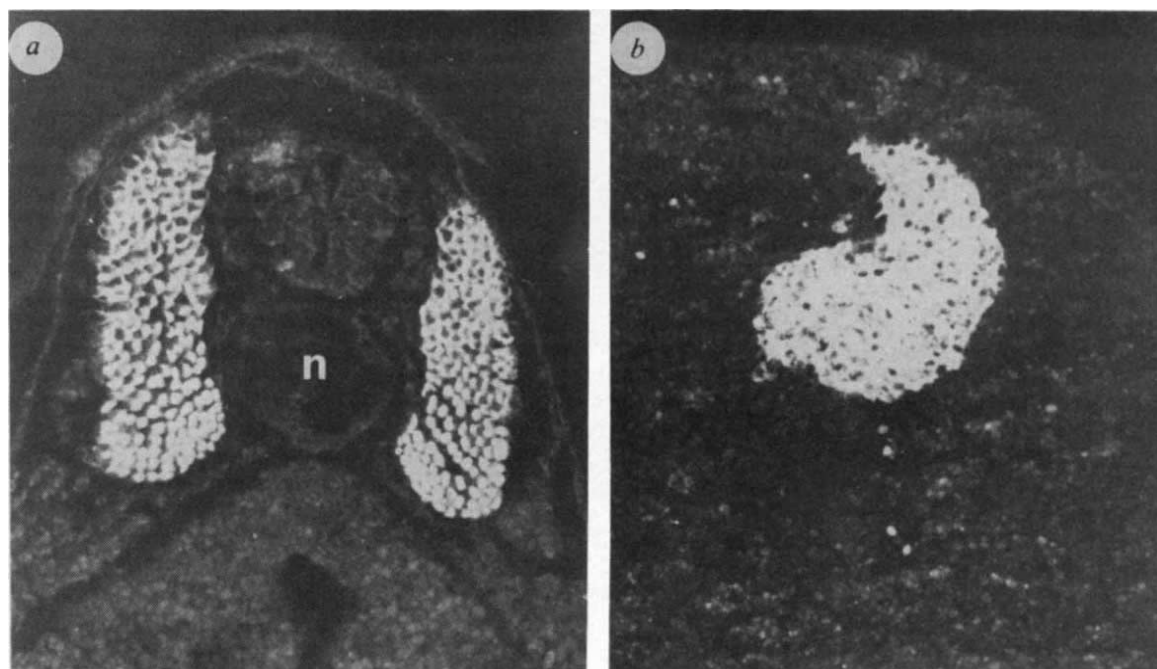


Fig.1 Muscle cells formed by embryonic induction in early amphibian development are arranged in homogeneous coherent groups sharply demarcated from adjacent non-muscle cells. *a*, Transverse section through a stage-36 tadpole, showing blocks of muscle cells in white (fluorescent antibody)¹¹. Muscle is derived embryologically from the same sheet of tissue (mesoderm) as the notochord (*n*). *b*, Transverse section through the animal cell region of a conjugate assembled from animal and vegetal tissues of a blastula, and cultured until stage 28. The cells which are positive for the muscle-specific antibody (seen white) are in a single coherent group.

so. This result is most simply interpreted as a community effect, in which the ability of a cell to respond to induction by differentiating as muscle is greatly increased if it is surrounded by other cells responding in the same way at the same time.

One criticism of these experiments is that the main conclusion rests on the failure of monolayer cells to differentiate as muscle. Unfortunately there is a very limited range of cell-type-specific antibodies for tissues other than muscle in early *Xenopus* embryos, and many of these are not diagnostic at the single cell level. Therefore it has not been possible so far to show that the monolayer animal cells differentiate in any other identifiable way. For this reason it is particularly important to establish that the test animal cells were as viable in monolayer as in solid configuration. This has been checked in several ways. In one series of experiments, animal cells were unlabelled when placed in conjugates, and were subsequently recognized by other means. At the end of the incubation period, sandwich conjugates were injected, while still in a cytochalasin-containing gel, with either [³H]uridine or [³H]histidine. Three hours later, conjugates were fixed and processed as usual. Subsequent autoradiography

showed that, at the end of the culture period, animal cells were synthesizing RNA and protein to a similar extent, whether they were in monolayer or solid reaggregate configuration. There was therefore no difference in synthetic activity between cells according to whether they differentiated as muscle or not. Another series of experiments involved mixing animal cells which had been treated exactly as in the monolayer experiments above with an excess of equatorial cells from several blastulae. Subsequent analysis showed that animal cells surrounded by equatorial cells that were positive for the muscle antigen were themselves also positive. This shows that animal cells, treated as in monolayer experiments, are viable and capable of differentiation if the surrounding cellular environment is appropriate. Details of these controls will be published elsewhere.²⁶

The mode of action of the community effect is not known, but may be envisaged as an autocrine or paracrine effect like that in tumour cells¹³ and blood cells^{14,15}. Perhaps vegetally induced animal cells release unstable or slowly diffusing substances, which act as differentiation factors, and which enhance muscle gene activation in proportion to their concentration; this would be much higher in solid reaggregates than in monolayer cells. The community effect described here bears a superficial resemblance to other embryological phenomena, such as homeogenetic induction¹⁶, mass effects^{17,18}, and feedback differentiation effects^{19,20}, but clearly differs from these in ways to be explained elsewhere. The community effect is consistent with results of Symes *et al.*²¹ who find that a soluble inducer (XTC factor) fails to induce dispersed animal cells to form muscle, but does so successfully in similar cells arranged as a solid tissue. The results reported here are also consistent with the blastula cell transfer experiments of Heasman *et al.*²².

The community effect, if confirmed in further work, may be of general significance as a morphogenetic mechanism. Embryonic induction is of major importance in vertebrate development, and is also involved in the early development of many invertebrates^{23,24}. Whether a particular animal cell will respond

Table 1 Muscle gene activation by embryonic induction is enhanced by a community effect.

Arrangement of animal cells in vegetal sandwich	Total number cells scored*	Normal nuclei (Hoechst)†	Cells positive for muscle antibody‡
Monolayer reaggregate	281	79%	0%
Solid reaggregate	2,390	83%	64%

* Analyses of four monolayers and four solid reaggregates. Only those cells were scored which were of animal cell origin (radioactively labelled). † Condensed nuclei were excluded; expressed as per cent of total cells. ‡ Only cells of animal origin with normal nuclei (Hoechst) are included in this column.

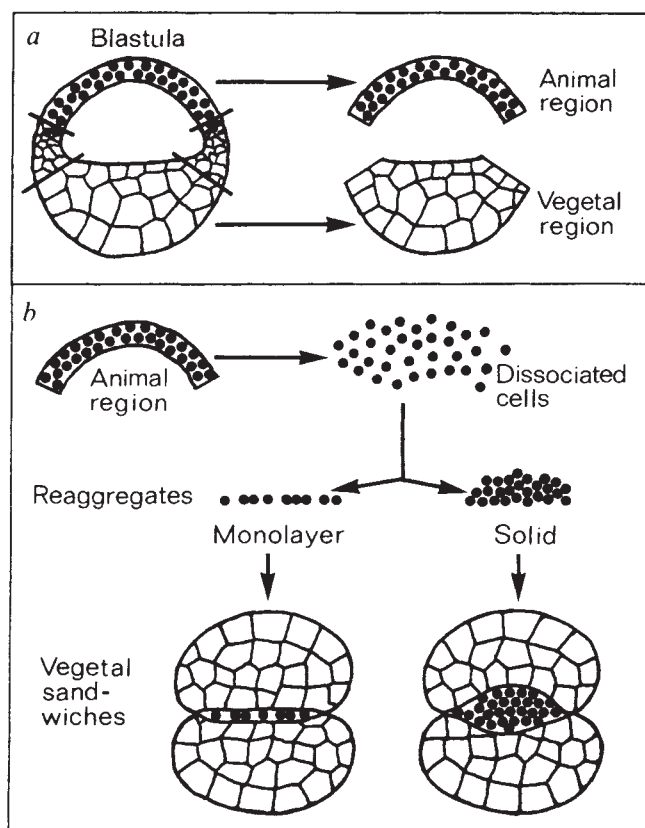


Fig. 2 Diagrams illustrating the procedures used to place animal cells of a blastula into vegetal sandwiches in either monolayer or solid configurations. Vegetal cells are indicated by their outlines as large cuboidal cells; animal cells are shown diagrammatically as black circles. *a*, A blastula is cut into animal and vegetal regions, discarding the equatorial cells. *b*, Animal regions (radioactively labelled) are dissociated in Ca^{2+} -free medium into their component cells. These are reagggregated in Ca^{2+} -containing medium into monolayer or solid configurations, and then incorporated in between two undissociated vegetal pieces of unlabelled blastulae. The vegetal sandwiches are cultured until stage 28, when they are fixed, sectioned and processed as described in the text.

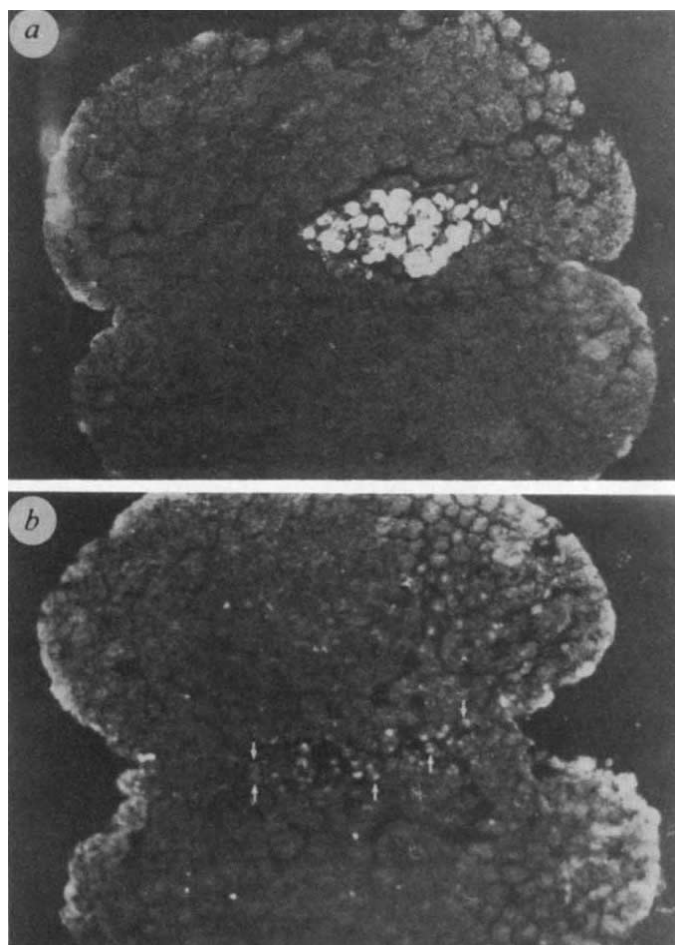


Fig. 3 Transverse sections of vegetal sandwiches containing animal region cells arranged as a solid group or as a monolayer. *a*, Animal cells reagggregated as a group before implantation into the sandwich. *b*, Similar cells reagggregated as a monolayer (white arrows). Muscle gene expression (white cells due to a fluorescent muscle-specific antibody¹¹) is seen in the solid group (*a*) but not in the monolayer (*b*). Animal region cells (radioactively labelled) were identified by autoradiography after sections had been reacted with antibody.

to muscle-forming induction in amphibians is determined rather imprecisely by how far the vegetal inducer substances spread within the time available¹². The community effect would help to ensure that a cell in the borderline area, the floating voter, is encouraged more strongly to conform to one group of neighbours all of which respond to induction, or to another group all of which fail to respond through lack of inducer. A mechanism of this kind might occur widely in development because solid blocks of tissue are commonly formed by induction from continuous sheets of cells. The community effect would not

contribute to cells like the neural crest which are widely dispersed²⁵, nor to others such as epidermis where single cells can differentiate autonomously^{8,9}. Indeed, most aspects of animal development seem to proceed by the cooperation of several contributory processes, no one of which is individually indispensable; the community effect probably falls into this class.

I thank J. Brockes for antibody, W. Westley and A. Pluck for histological help, Nick Hopwood for comments on the manuscript, and the Cancer Research Campaign for support.

Received 29 July; accepted 21 November 1988.

1. Saxen, L. & Toivonen, S. *Primary Embryonic Induction* (Academic, New York, 1962).
2. Kratochwil, K. in *Cell Interactions and Development: Molecular Mechanisms* (ed. Yamada, K. M.) 99-122 (Wiley, New York, 1983).
3. Gurdon, J. B. *Development* **99**, 285-306 (1987).
4. Nieuwkoop, P. D. *Wilhelm Roux Arch.* **162**, 341-373 (1969).
5. Nakamura, O., Takasaki, H. & Ishihara, M. *Int. Proc. Jap. Acad.* **47**, 313-318 (1971).
6. Gurdon, J. B., Fairman, S., Mohun, T. J. & Brennan, S. *Cell* **41**, 913-922 (1985).
7. Jonas, E. J., Sargent, T. D. & Dawid, I. B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5413-5417 (1985).
8. Jones, E. A. & Woodland, H. R. *Cell* **44**, 345-355 (1986).
9. Sargent, T. D., Jamrich, M. & Dawid, I. B. *Dev. Biol.* **114**, 238-246 (1986).
10. Smith, J. C. & Malacinski, G. M. *Dev. Biol.* **98**, 250-254 (1983).
11. Kintner, C. R. & Brockes, J. P. *Nature* **308**, 67-96 (1984).

12. Gurdon, J. B. *Development* (in the press).
13. Sporn, M. B. & Roberts, A. B. *Nature* **313**, 745-747 (1985).
14. Metcalf, D. *Proc. R. Soc. B* **230**, 387-423 (1987).
15. Farrar, J. J. *et al. Immun. Rev.* **63**, 129-166 (1982).
16. Mangold, O. & Spemann, H. *Wilhelm Roux Arch. EntwMech. Org.* **111**, 341-422 (1927).
17. Muchmore, W. B. *J. exp. Zool.* **134**, 293-313 (1957).
18. Grunz, H. *Wilhelm Roux Arch. dev. Biol.* **187**, 49-57 (1979).
19. Nevo, Z. & Dorfman, A. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2069-2072 (1972).
20. Meier, S. and Hay, E. D. *Dev. Biol.* **38**, 249-270 (1974).
21. Symes, K., Yaqoob, M. & Smith, J. C. *Development* (in the press).
22. Heasman, J., Wylie, C. C., Hausen, P. & Smith, J. C. *Cell* **37**, 185-194 (1984).
23. Priess, J. R. & Thomson, J. N. *Cell* **48**, 241-250 (1987).
24. Martinez-Arias, A., Baker, N. E. & Ingham, P. W. *Development* **103**, 157-170 (1988).
25. Le Douarin, N. *Cell* **38**, 353-360 (1984).
26. Gurdon, J. B. *Development* (submitted).

Chemotropic guidance of developing axons in the mammalian central nervous system

Marc Tessier-Lavigne*†, Marysia Placzek*‡, Andrew G. S. Lumsden§, Jane Dodd*‡ & Thomas M. Jessell*†

* Center for Neurobiology and Behavior, † Howard Hughes Medical Institute and Department of Biochemistry and Molecular Biophysics, and ‡ Department of Physiology and Cellular Biophysics, Columbia University, New York 10032, USA § Department of Anatomy, Guy's Hospital Medical School, London, UK

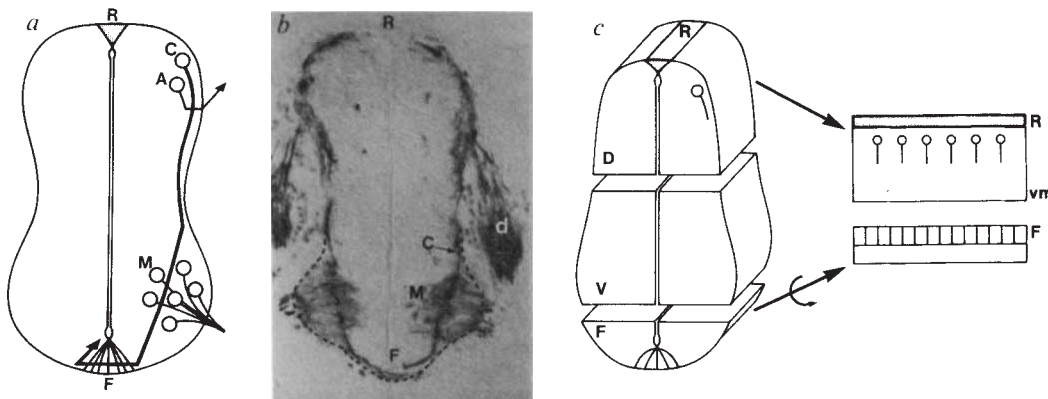
In the developing nervous system, axons project considerable distances along stereotyped pathways to reach their targets. Axon guidance depends partly on the recognition of cell-surface and extracellular matrix cues derived from cells along the pathways¹. It has also been proposed that neuronal growth cones are guided by gradients of chemoattractant molecules emanating from their intermediate or final cellular targets². Although there is evidence that the axons of some peripheral neurons in vertebrates are guided by chemotropism^{3,4} and the directed growth of some central axons to their targets is consistent with such a mechanism⁵⁻⁷, it remains to be determined whether chemotropism operates in the central nervous system. During development of the spinal cord, commissural axons are deflected towards a specialized set of midline neural epithelial cells, termed the floor plate⁸⁻¹⁰, which could reflect guidance by substrate cues or by diffusible chemoattractant molecules. Here we provide evidence in support of chemotropic guidance by demonstrating that the rat floor-plate cells secrete a

diffusible factor(s) that influences the pattern and orientation of commissural axon growth *in vitro* without affecting other embryonic spinal cord axons. These findings support the hypothesis² that chemotropic mechanisms guide developing axons to their intermediate targets in the vertebrate CNS.

The existence of a floor plate-derived chemotropic factor was established by examining the effect of floor-plate explants and defined medium conditioned *in vitro* by the floor plate on axon outgrowth from explants of dorsal spinal cord. Dorsal explants were taken from rat embryos at embryonic day (E) 11, an age at which commissural neurons are beginning to differentiate and extend processes⁸ (Fig. 1a and b). Explants were cultured in a collagen gel matrix (Fig. 1c) capable of stabilizing gradients of diffusible molecules¹¹. In control experiments, there was no outgrowth from 75% of dorsal explants cultured alone for 39-44 h, and only sparse outgrowth from the remainder (Fig. 2a and Table 1, *n* = 107). In contrast, dorsal explants cultured with an E11 floor-plate explant placed 100-400 µm from their ventral-most edge (Fig. 1c) showed clear axon outgrowth within 20-24 h (not shown). After 39-44 h, a characteristic pattern of outgrowth was observed from all of these explants, with most axons projecting to the floor plate from their ventral-most edge in thick fascicles (Fig. 2b and Table 1, *n* = 116 co-cultures). The axons that projected into the collagen gel were not preceded or accompanied by migrating cells, as assessed by staining with the nuclear dye, Hoechst 33258 (*n* = 8). Marked outgrowth was also observed from dorsal explants cultured alone but exposed to medium conditioned by E11 floor plate (Table 1), but in this case axons emerged from all edges of the explant (M.P. *et al*, manuscript in preparation). These experiments show that the floor plate secretes a diffusible factor that promotes axon outgrowth from dorsal explants.

The axons that projected from dorsal explants in the presence

Fig. 1 a and b, Trajectory of commissural axons in the embryonic rat spinal cord. a, Schematic diagram of a transverse section of an E12 rat spinal cord (E0, day of vaginal plug) showing the location of the first three classes of differentiated neurons and their prospective axonal trajectories. Motoneurons (M) differentiate in the ventral region of the spinal cord from E10.5 (refs 8, 20) and extend axons to their target muscles. Commissural (C) and association (A) neurons differentiate



from E11 (refs 8, 20) in the dorsal region of the spinal cord, adjacent to the roof plate¹⁰ (R). Association axons project laterally to join the ipsilateral lateral funiculus^{8,20} (arrow). Commissural axons grow ventrally along the lateral margin of the spinal cord to the motor column, then alter their trajectory and course directly through the nascent motor column to the floor plate (F). Because these axons appear to home in on the floor plate through a cellular environment that does not contain a preformed pathway^{6,9}, it seems possible that they are guided by chemoattractant molecules secreted by the floor plate. After crossing the midline of the spinal cord at the floor plate, commissural axons turn by 90° to form longitudinal projections in the contralateral ventrolateral funiculus²⁰ (arrow). b, Transverse section of an E12 rat spinal cord at the cervical level, stained by the immunoperoxidase method with a monoclonal antibody (4D7) (ref. 21) to the TAG-1 antigen (see ref. 10 for methods). Motor and commissural axons, but not association axons, express TAG-1 (ref. 10). Commissural axons project towards the floor plate (arrow). TAG-1⁺ neurons in the dorsal root ganglia^{10,21} (d) can also be seen. Note that the spinal cord shown here is at a more advanced stage of development than those used for explant cultures (see below). Calibration, 100 µm. c, Schematic diagram of the main experimental protocol. Each section of E11 spinal cord (see below) was dissected into three portions: a 'dorsal explant' (D) comprising the dorsal third or half, a 'floor-plate explant' (F) comprising the ventral-most fifth, and the remaining 'ventral explant without floor plate' (V) which contained most of the motoneurons. Explants were then embedded in three-dimensional collagen matrices in appropriate orientations. For instance, dorsal and floor-plate explants were co-cultured as depicted on the right: the dorsal explant was placed with its ventral-most edge (vm) facing the ventral-most edge of the floor-plate explant.

Methods. Spinal cord explants were taken from rat embryos at the 24-27 somite stage (E11-E11.5). Embryos were placed in L15 medium (Gibco), and tungsten needles were used to remove a two-segment portion of the vertebral column 14 segments from the most recently segmented somites. These pieces of tissue were incubated with Dispase (Boehringer, 1 mg ml⁻¹ in L15) for 30 min at room temperature, which facilitated excision of the spinal cord. After further dissection, the explants were embedded in a three-dimensional collagen gel matrix as described³, and cultured in supplemented Ham's F12 medium²² with 5% horse serum (Gibco) at 37 °C in a 5% CO₂ environment.

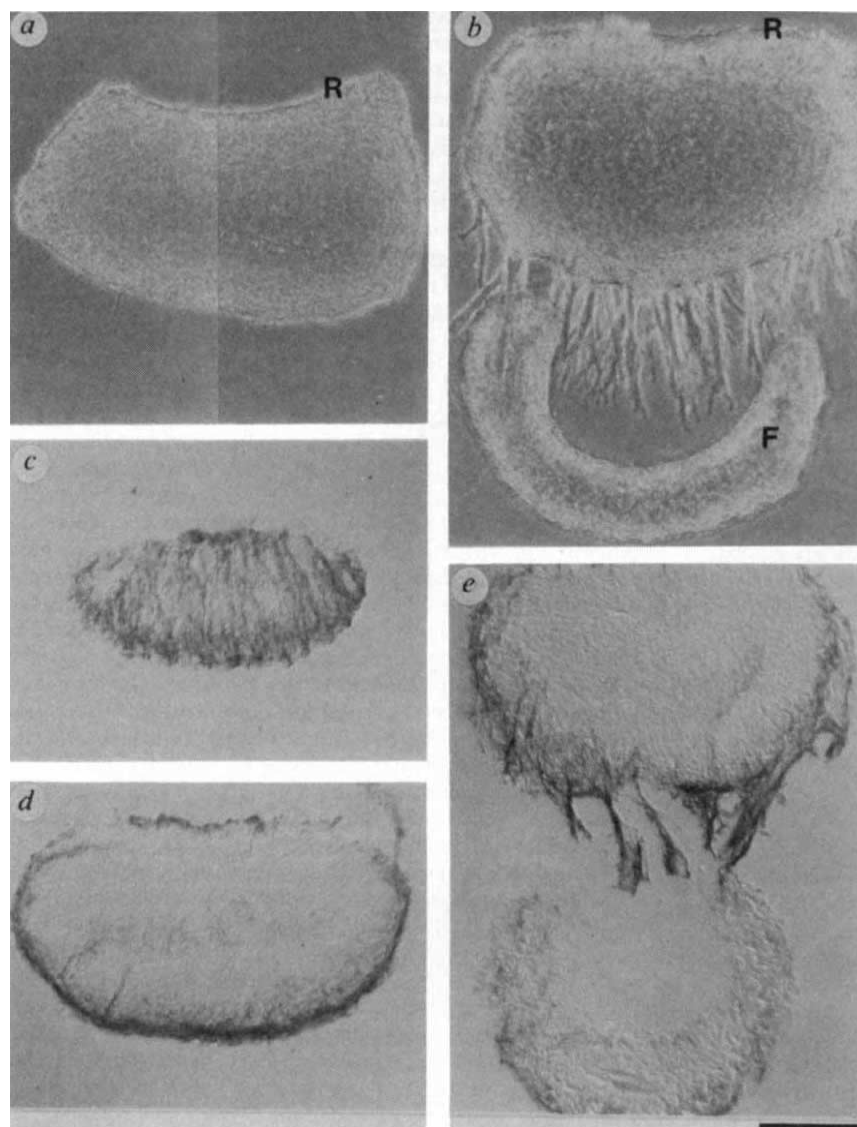


Fig. 2 The floor plate promotes commissural axon outgrowth from dorsal spinal cord explants. All explants were taken from E11 rat embryos and cultured in collagen gels, as described in Fig. 1c. In *a-e*, the dorso-ventral axis of dorsal explants is oriented vertically, ventral downwards. *a*, Little axon outgrowth is observed from a dorsal spinal cord explant cultured alone for 44 h. Abbreviations as in Fig. 1. *b*, Extensive axon outgrowth is apparent from a dorsal explant cultured for 44 h with a floor-plate explant. *c* and *d*, TAG-1 expression observed by immunoperoxidase labelling in cryostat sections through two different levels of a dorsal explant cultured without floor plate for 44 h. Extensive growth of TAG-1⁺ axons, presumably from commissural neurons^{10,12}, is observed along a dorso-ventral trajectory (*c*). These axons then extend along the inside perimeter of the explant (*d*), and do not project into the collagen gel. *e*, TAG-1 expression observed in a cryostat section of a dorsal explant cultured with floor plate for 40 h. TAG-1⁺ axons are oriented along a dorso-ventral trajectory (not shown), and project into the collagen gel. Note that the floor-plate explant contains a few TAG-1⁺ motor axons. Calibrations: *a*, *b*, 150 μ m; *c*, *d*, 135 μ m; *e*, 130 μ m. R and F as in Fig. 1.

Methods. For immunostaining, explants were fixed after 39–44 h in culture by immersing the collagen gels in 0.12 M phosphate buffer (pH 7.4) containing 4% paraformaldehyde for 3 h at 4 °C, then transferred to 30% sucrose in 0.1 M phosphate buffer for 48–72 h at 4 °C. A piece (~0.6 cm $\times$ 0.6 cm) of each gel containing the explants was then cut out and frozen in OCT (TissueTek) compound, and 15–20 μ m cryostat sections were collected on to gelatin-subbed slides. Staining was as described¹⁰.

of floor plate appear to derive from commissural neurons as they express the TAG-1 glycoprotein (Fig. 2*e*, $n = 51$ explants sectioned). Studies both *in vivo* and *in vitro* indicate that TAG-1 is selectively expressed by commissural neurons in the dorsal spinal cord (refs 10 and 12; Fig. 1*b*). Many TAG-1⁺ axons were also observed within control explants (Fig. 2*c* and *d*, $n = 16$ explants sectioned) and were oriented along the original dorso-ventral axis of the explant (Fig. 2*c*). Instead of projecting into the collagen gel, however, these axons remained confined to the explant, extending along its inside perimeter (Fig. 2*d*). The floor plate-derived factor therefore appears to promote commissural axon outgrowth from the neural epithelium into the collagen gel without detectable effect on the differentiation or survival of commissural neurons. By 60 h in culture, many axons had projected from control explants (15.5 ± 1.8 (s.e.) axon bundles per explant, $n = 11$), which could suggest that the effect of the floor plate is simply to increase the rate of commissural axon growth. Several lines of evidence argue against this interpretation, however. First, many of the axons that emerged by 60 h were TAG-1⁻ (not shown), and may therefore have derived from association neurons (see below). Second, most of these axons emerged from the sides and not the ventral-most edge of dorsal explants (not shown), unlike the projection pattern observed in the presence of floor plate. Finally, the density and length of TAG-1⁺ axons in dorsal explants cultured for 20 h (that is, a time at which axons are beginning to project from

dorsal explants cultured with floor plate) appeared to be similar in the presence ($n = 8$) and absence ($n = 10$) of floor plate, showing that the floor plate does not markedly enhance the initial rate of commissural axon growth.

The ability to evoke commissural axon outgrowth from E11 dorsal explants was confined largely to the floor plate over the period in which commissural axons project to the floor plate *in vivo* (that is E11–14)⁸. Explants of E13–14 floor plate had the same effect as E11 floor plate in evoking profuse outgrowth of TAG-1⁺ axons from dorsal explants (Table 1). This effect was not mimicked by explants of E11–14 dorsal spinal cord ($n = 27$) or E11 ventral spinal cord ($n = 16$), but E13–14 ventral spinal cord did evoke the outgrowth of a small number of TAG-1⁺ axons above background (Table 1). It is possible that ventral spinal cord cells at this age secrete low amounts of a factor that affects commissural axon outgrowth. Alternatively, the factor secreted by the floor plate may bind to the ventral spinal cord *in vivo* and be released slowly by ventral explants *in vitro*.

The effect of the floor plate on axon outgrowth appears to be selective for commissural neurons. Neurofilament-reactive axons which did not express TAG-1 and presumably derived from association neurons (ref. 10 and Fig. 1), were also present within dorsal explants cultured for 39–44 h with floor plate (not shown), but they were not observed projecting from the explants towards the floor plate ($n = 10$). Also, motor axons extended into the collagen gel from ventral explants grown alone (not

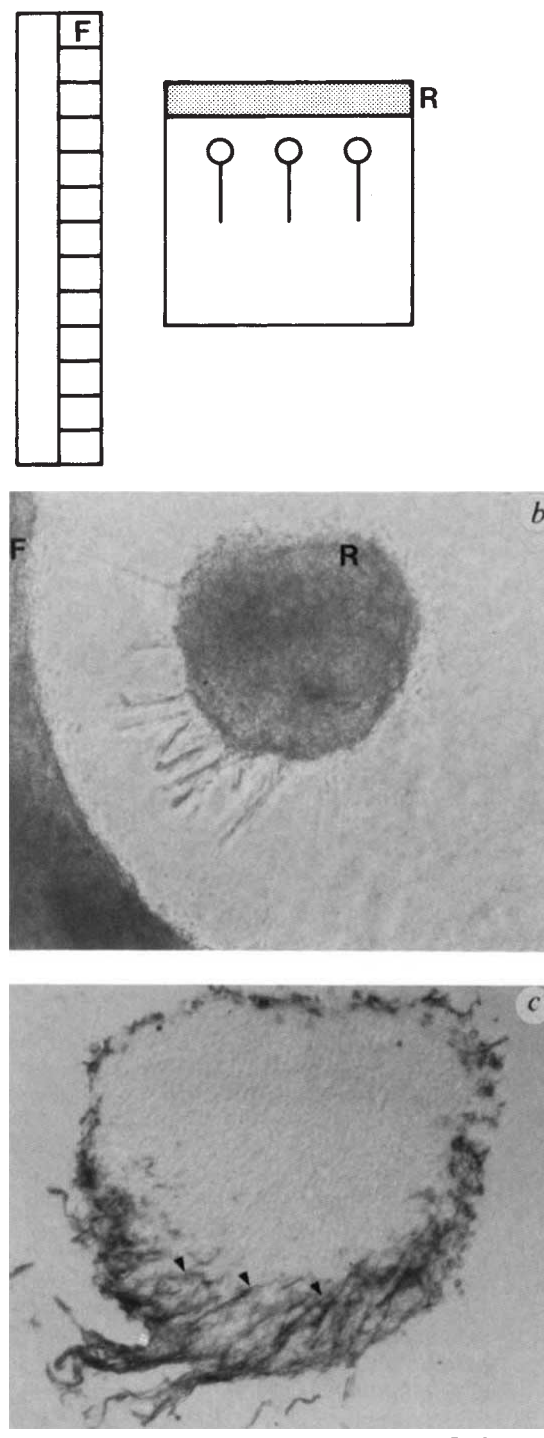


Fig. 3. The floor plate orients the growth of commissural axons *in vitro*. *a*, Schematic diagram of the experimental protocol. An E11 dorsal explant was embedded in a collagen gel with its roof plate placed at right angles to a long strip of E11 floor plate. To prevent the floor plate from curling up extensively (see Fig. 2*b* and *e*), the ventral half of the spinal cord was left attached to the floor plate. *b*, After 42 h, most of the axons projecting from a dorsal explant emerged near to, and were oriented towards, the floor plate. In 30 similar co-cultures, the majority of the emerging axons were oriented towards the floor plate. This orienting effect was observed whether or not the floor-plate explant curled. *c*, TAG-1 immunostaining of a cryostat section from a similar explant to that shown in (*b*). TAG-1⁺ axons are oriented towards the floor plate within the dorsal explant itself (arrowheads). This behaviour was observed in the 14 explants cultured with floor plate that were sectioned, but not in explants cultured with E11 dorsal (*n* = 4) or E11 ventral (*n* = 6) spinal cord. Calibrations: *b*, 120 μ m; *c*, 50 μ m.

a

Table 1 Characteristics of axon bundles emerging from E11 dorsal spinal cord explants

Condition	Number of explants	Number of bundles per explant	
		Mean ($\pm$ s.e.)	Bundle length (μ m) Mean ($\pm$ s.e.)
Dorsal explant (D)	107	0.4 $\pm$ 0.1	63 $\pm$ 3.4
D + E11 floor plate	64	17 $\pm$ 1.1	129 $\pm$ 5.4
D + E11 floor plate-conditioned medium	13	25 $\pm$ 2.5	126 $\pm$ 11
D + E13-14 floor plate	13	29 $\pm$ 3.5	154 $\pm$ 12
D + E13-14 ventral spinal cord	27	3.2 $\pm$ 0.6	88 $\pm$ 8.3

Dorsal explants were cultured for 39–44 h under the following conditions: alone (control), 100–400 μ m from different explants as indicated, or with E11 floor plate-conditioned medium (see below). Measurements were made from camera-lucida tracings or photographs of the explants. The number and length of bundles emerging from dorsal explants cultured with E11 or E13–14 floor plate explants, or with E11 floor plate-conditioned medium were significantly greater than those in control explants ($P < 0.0005$, Student's *t* test). E13–14 ventral spinal cord evoked a small but statistically significant amount of axon outgrowth above background ($P < 0.0005$), but this effect was significantly smaller than that observed with floor plate or floor plate-conditioned medium ($P < 0.0005$). Moreover, as shown here, the bundles that emerged in these cases were considerably shorter than those observed with floor plate ($P < 0.0005$), and were thinner (not shown). The number and length of bundles emerging from dorsal explants cultured with E11–14 dorsal spinal cord, E11 ventral spinal cord, or medium conditioned by the remainder of the spinal cord were not significantly different from those observed in control explants ($P > 0.1$). For conditioned-medium experiments, culture medium (Fig. 1 legend) was exposed for 48 h to E11 floor plate tissue or to the remainder of the E11 spinal cord (M.P. *et al.*, manuscript in preparation). In both cases, tissue from 100 E11 embryos was pooled to condition 400 μ l medium.

shown), but the extent and pattern of motor axon outgrowth was not altered when they were cultured with floor plate ($n = 7$).

These findings show that the floor plate promotes commissural axon outgrowth, but not whether it can orient the growth of these axons. To test for a chemotropic action, dorsal explants were cultured adjacent to an E11 floor-plate explant placed parallel to the original dorso-ventral axis of the spinal cord (Fig. 3*a*). Most axons that emerged from dorsal explants projected towards the floor plate (Fig. 3*b*, $n = 30$). Moreover, TAG-1⁺ axons within the dorsal explants themselves were clearly oriented towards the floor plate (Fig. 3*c*, $n = 14$ explants sectioned), and were not oriented in this way when E11 dorsal ($n = 4$) or ventral ($n = 6$) explants were substituted for floor plate. The floor plate thus causes commissural axons to deviate from their expected trajectory, consistent with a chemotropic action of the floor plate-derived factor. In particular, these observations show that the floor plate can orient the growth of commissural axons within a piece of explanted neural epithelium, which presumably approximates the *in vivo* environment of the axons more closely than a collagen gel.

The identity of the factor(s) secreted by the floor plate is unknown. The effect of the floor plate on commissural axon outgrowth was not mimicked by 7.5S nerve growth factor (100 ng ml⁻¹, $n = 16$), which can exert a chemotropic action on peripheral neurites *in vitro*^{13,14}, or by EHS-laminin (100 μ g ml⁻¹, added to the collagen gel and the culture medium, $n = 17$), which promotes neurite outgrowth from many central and peripheral neurons^{15–17}.

The idea that axons are guided by a series of tropic factors released by their successive intermediate targets was first proposed in 1892 (ref. 2), but it has not been generally accepted¹⁸. Our findings have established that a defined intermediate target, the floor plate, secretes a diffusible factor(s) that functions as a selective chemoattractant for embryonic commissural axons

in vitro. Chemotropism may be only one of several guidance mechanisms that operate successively or coordinately to define the trajectory of commissural axons *in vivo*. Their initial ventral trajectory through the dorsal spinal cord (see Figs 1 and 2c) may be determined by extracellular matrix cues, in particular laminin (ref. 19; J.D. *et al.*, manuscript in preparation). Moreover, once commissural growth cones arrive at the midline, their subsequent guidance may be regulated by contact-dependent interactions with floor-plate cells¹⁰. Identification of the floor plate-derived chemotropic factor should permit experimental analysis of the contribution of chemotropism to the guidance of commissural axons *in vivo*. More generally, our observations raise the possibility that chemotropism represents a major mechanism of axon guidance in both the central and peripheral nervous systems.

We thank Richard Axel, Alun Davies and our colleagues for critical comments on the manuscript. Supported by the NIH, the McKnight Foundation, the Irma T. Hirsch Charitable Trust, and a NATO twinning grant. T.M.J. is an investigator of the Howard Hughes Medical Institute.

Received 14 September; accepted 4 November 1988.

- Jessell, T. M. *Neuron* **1**, 3–13 (1988).
- Ramon y Cajal, S. *Histologie du Systeme Nerveux de l'Homme et des Vertebres* Vol. 1 657–664 (Consejo superior de investigaciones cientificas, Madrid, 1909).
- Lumsden, A. G. S. & Davies, A. M. *Nature* **306**, 786–788 (1983).
- Lumsden, A. G. S. & Davies, A. M. *Nature* **323**, 538–539 (1986).
- Singer, M., Nordlander, R. H. & Egar, M. J. *comp. Neurol.* **185**, 1–22 (1979).
- Scott, T. M. & Bunt, S. M. *J. Embryol. exp. Morph.* **91**, 181–195 (1986).
- Harris, W. A. *Nature* **320**, 266–270 (1986).
- Altman, J. & Bayer, S. A. *Adv. anat. Cell Biol.* **85**, 1–164 (1984).
- Holley, J. A. & Silver, J. *Dev. Biol.* **123**, 375–388 (1987).
- Dodd, J., Morton, S. B., Karagogeos, D., Yamamoto, M. & Jessell, T. *Neuron* **1**, 105–116 (1988).
- Ebendal, T. & Jacobson, C. O. *Expl Cell Res.* **105**, 379–387 (1977).
- Karagogeos, D., Dodd, J. & Jessell, T. M. *Soc. Neurosci. Abstr.* **14**, 108.7 (1988).
- Letourneau, P. C. *Dev. Biol.* **66**, 183–196 (1978).
- Gundersen, R. W. & Barrett, J. N. *Science* **206**, 1079–1080 (1979).
- Rogers, S. L., Letourneau, P. C., Palm, S. L., McCarthy, J. & Furcht, L. T. *Dev. Biol.* **98**, 212–220 (1983).
- Lander, A. D., Fujii, D. K. & Reichardt, L. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2183–2187 (1985).
- Hall, D. E., Neugebauer, K. M. & Reichardt, L. F. *J. Cell Biol.* **104**, 623–634 (1987).
- Hamburger, V. J. *Neurosci.* **8**, 3535–3540 (1988).
- Letourneau, P. C., Madsen, A. M., Palm, S. L. & Furcht, L. T. *Dev. Biol.* **125**, 135–144 (1988).
- Wentworth, L. E. J. *comp. Neurol.* **222**, 96–115 (1984).
- Yamamoto, M., Boyer, A. M., Crandall, J. E., Edwards, M. & Tanaka, H. J. *Neurosci.* **6**, 3576–3594 (1986).
- Romijn, H. J., Gabets, A. M. M. C., Mud, M. T. & Walters, P. S. *Dev. Brain Res.* **2**, 583–589 (1982).

A malaria T-cell epitope recognized in association with most mouse and human MHC class II molecules

F. Sinigaglia, M. Guttinger, J. Kilgus, D. M. Doran, H. Matile, H. Etlinger, A. Trzeciak, D. Gillessen & J. R. L. Pink

Central Research Units, F. Hoffmann-La Roche & Co. Ltd, 4002 Basel, Switzerland

An ideal vaccine should elicit a long lasting immune response against the natural parasite, both at the T- and B-cell level. The immune response should occur in all individuals and be directed against determinants that do not vary in the natural parasite population. A major problem in designing synthetic peptide vaccines is that T cells generally recognize peptide antigens only in association with one or a few of the many variants of major histocompatibility complex (MHC) antigens^{1,2}. During the characterization of epitopes of the malaria parasite *Plasmodium falciparum* that are recognized by human T cells, we analysed a sequence of the circumsporozoite protein, and found that synthetic peptides corresponding to this sequence are recognized by T cells

in association with many different MHC class II molecules, both in mouse and in man. This region of the circumsporozoite protein is invariant in different parasite isolates^{3,4}. Peptides derived from this region should be capable of inducing T-cell responses in individuals of most HLA-DR types, and may represent good candidates for inclusion in an effective anti-malaria peptide vaccine.

We had observed that peripheral blood mononuclear cells (PBMC) from the majority of immune and non-immune donors responded *in vitro* to a peptide, CS.T3, corresponding, with the exception of two cysteine-alanine substitutions, to residues 378–398 of the circumsporozoite (CS) protein⁵. This suggested that the peptide could associate with many different MHC class II molecules. To test this hypothesis, cells from 20 non-immune donors (all with titres of anti-sporozoite antibodies <1:80) were challenged *in vitro* with CS.T3 at 10 μ M. We derived a total of 298 CS.T3-specific T-cell clones from eight different donors. None of the clones proliferated in the presence of a control peptide (CS protein 325–341, data not shown).

The MHC restriction of CS.T3-specific T-cell clones was assessed using anti-MHC-class II monoclonal antibodies (mAbs). The proliferation of all 187 clones tested was inhibited by the mAb E.31 (ref. 6), which recognizes a monomorphic HLA-DR determinant. Neither anti-DP⁷ nor anti-DQ⁸ mAbs inhibited T-cell proliferation, and fibroblasts transfected with the DR α and DR β genes presented antigen to selected clones, confirming (Table 1) that the DR molecule is the restriction element for CS.T3-specific T-cell clones. The DR restriction specificity of each T-cell clone was then determined using a panel of HLA-DR-homozygous antigen-presenting cells. As shown in Table 1, T-cell clones responded equally well to CS.T3 when presented on autologous Epstein-Barr virus-transformed B (EBV-B) cells or on an EBV-B line homozygous for one of the donor's DR specificities. The peptide CS.T3 was recognized in association with at least seven different DR molecules (DR1, 2, 4, 5, w6, 7 and 9).

To determine whether T-cell clones restricted by different DR molecules recognize different determinants on CS.T3, we assayed proliferative responses to a series of peptides truncated at either the N or C terminus of the CS.T3 sequence (Table 2). Peptides 380–398, 378–395 and larger peptides were stimulatory for all T-cell clones examined. However, when shorter peptides were tested, distinct recognition patterns emerged. At the two extremes stand the DR2- and DR5-restricted clones. The minimal regions that are stimulatory for DR2- and DR5-restricted clones correspond to residues 385–395 and 381–390 respectively. The minimal stimulatory region for DR4-restricted clones lies between residues 383 and 394, the regions for DRw6 and DR7 correspond to residues 381–392 and 381–394 respectively, and for DR1 to residues 382–395 (data not shown) or 381–392. DR9-restricted T-cell clones (four out of four tested) did not recognize peptides 382–398 and 381–398 over a wide range of concentrations. Further removal of Lys 382 and Ile 383, however led to reappearance of recognition. The still shorter peptide 385–398 was not stimulatory at all (Table 2). In conclusion, closely overlapping epitopes between residues 381–395 are recognized in slightly different ways in association with different DR antigens, consistent with earlier studies which showed a correlation between the MHC class II restriction element and the fine specificity of antigen recognition by mouse T cells^{9–11}.

To see whether CS.T3 could induce helper T-cell function *in vivo*, we coupled the peptide to (NANP)₃; the repetitive NANP (Asn-Ala-Asn-Pro) sequence in the central domain of the CS protein being the major epitope recognized by anti-sporozoite antibodies^{12–14}. The (NANP)₃-CS.T3 peptide was administered to seven different mouse strains, and anti-(NANP)_n and anti-sporozoite antibody responses were measured (Table 3). All the strains mounted an antibody response against both (NANP)_n and sporozoites. Although C57BL/6 T cells can recognize (NANP)_n, T cells from the other strains do not respond to the

Table 1 Restriction specificity of CS.T3-specific T-cell clones

APC	DP52	MG30	DP11	SD13	BR82	SD22	MG15
Autologous	<u>14,616</u>	<u>10,241</u>	<u>10,387</u>	<u>29,031</u>	<u>22,355</u>	<u>29,889</u>	<u>10,594</u>
Autologous + anti-DR	<u>1,021</u>	<u>307</u>	<u>463</u>	<u>1,522</u>	<u>1,257</u>	<u>729</u>	<u>1,100</u>
Autologous + anti-DQ	<u>17,296</u>	<u>10,331</u>	<u>12,609</u>	<u>22,448</u>	<u>23,012</u>	<u>25,565</u>	<u>13,407</u>
Autologous + anti-DP	<u>14,210</u>	<u>9,724</u>	<u>10,655</u>	<u>21,304</u>	<u>22,462</u>	<u>30,248</u>	<u>10,315</u>
DR1 (EDR)	<u>38,226</u>	<u>441</u>	<u>913</u>	<u>810</u>	<u>1,100</u>	<u>404</u>	<u>977</u>
DR2 (NOL)	<u>514</u>	<u>34,388</u>	<u>754</u>	<u>1,024</u>	<u>941</u>	<u>585</u>	<u>609</u>
DR4 (BSM)	<u>724</u>	<u>394</u>	<u>29,210</u>	<u>991</u>	<u>882</u>	<u>354</u>	<u>570</u>
DR5 (ATH)	<u>301</u>	<u>474</u>	<u>925</u>	<u>50,400</u>	<u>1,152</u>	<u>1,610</u>	<u>796</u>
DRw6 (BEC)	<u>845</u>	<u>765</u>	<u>739</u>	<u>784</u>	<u>18,109</u>	<u>798</u>	<u>697</u>
DR7 (EKR)	<u>305</u>	<u>450</u>	<u>434</u>	<u>627</u>	<u>1,209</u>	<u>30,113</u>	<u>358</u>
DR9 (DKB)	<u>523</u>	<u>614</u>	<u>638</u>	<u>662</u>	<u>508</u>	<u>365</u>	<u>40,998</u>

Peripheral blood mononuclear cells (PBMC) from eight HLA-DR-typed European donors: MG(DR2, 9), DP(DR1, 4), JK(DR5, 7), BR(DR4, w6), BH(DR1, 3), AH(DR1, 4), SD(DR5, 7), PE(DR5, w6) with no history of malaria infection were stimulated with peptide CS.T3 (10 μ M), expanded in IL-2-containing medium and cloned as previously described⁵. To test antigen reactivity and restriction specificity of the clones in a proliferative assay, cloned T cells (2×10^4) were co-cultured in triplicate with 10^4 antigen-presenting cells (APC) (irradiated autologous or DR-homozygous EBV-B cells) in 0.2 ml complete medium with CS.T3 (10 μ M) or without antigen. [³H]Thymidine incorporation was measured 72 h later. Results for representative clones are mean values of c.p.m. from triplicate cultures. Anti-HLA antibodies were added to cultures as a 1/100 dilution of ascites fluid. The numbers of clones with known restriction specificity from different donors and the numbers restricted to the DR alleles indicated, were as follows: MG: 12 clones (eight DR2, four DR9); DP: 11 clones (three DR1, eight DR4); JK: nine clones (nine DR5); BR: 10 clones (five DR4, five DRw6); BH: 16 clones (16 DR1); AH: eight clones (six DR1, two DR4); SD: 17 clones (four DR5, 13 DR7); PE: 13 clones (13 DR5). All the clones analysed were CD4⁺, CD8⁻. Four out of four tested DR2 restricted T-cell clones from donor MG (MG3, MG30, MG33 and MG36) could also be stimulated by the CS.T3 peptide presented by murine L cells transfected with DR α and DR2 β a, but not DR2 β 2b genes²³ (D. Altmann, D. Wilkinson and F. Sinigaglia, unpublished observations). Positive results are underlined.

repetitive sequence alone¹⁵⁻¹⁷. The fact that all strains tested responded to (NANP)₃ coupled to CS.T3 implies that the CS.T3 peptide is recognized in association with many different mouse as well as human MHC class II molecules.

Although there is evidence that some peptides can bind to more than one MHC allele², this is the first evidence for such permissive association. The CS.T3 peptide is recognized in association with at least seven different DR antigens (possibly more, as we lack information about the other DR alleles). From the known frequencies of the DR alleles in different populations¹⁸, we would expect that 80-90% of Caucasian or African individuals would carry at least one DR allele capable of being

recognized in association with CS.T3. The fact that cells from only eight of 20 non-immune individuals responded to CS.T3 *in vitro* is probably due to the low frequency of reactive T cells able to respond in our primary *in vitro* conditions, as most of the unresponsive individuals express DR alleles known to be presenting elements for CS.T3. The structural basis of this permissive association is unclear. As the DR α -chain shows very little polymorphism, it may be that CS.T3 binding is more influenced by the α - than the β -chain. Although CS.T3 was predicted to be a T-cell epitope¹⁹, some of the truncated peptides (for example 384-398) that are still recognized by T-cell clones are not so predicted. The CS.T3 sequence has a high amphipathic

Table 2 Definition of CS.T3 determinants recognized by T-cell clones restricted by different DR alleles

Peptide	Amino-acid sequence	DR1	DR2	DR4	Restriction DR5	DRw6	DR7	DR9
378-398	DIEKKIAKMEKASSVFNVNS	++	++	++	++	++	++	++
379-398	IEKKIAKMEKASSVFNVNS	++	++	++	++	++	++	++
380-398	EKKIAKMEKASSVFNVNS	++	++	++	++	++	++	++
381-398	KKIAKMEKASSVFNVNS	+	++	++	+	++	++	-
382-398	KIAKMEKASSVFNVNS	-	++	++	-	-	-	-
383-398	IAKMEKASSVFNVNS	-	++	+	-	-	-	+
384-398	AKMEKASSVFNVNS	-	++	-	-	-	-	++
385-398	KMEKASSVFNVNS	-	+	-	-	-	-	-
386-398	MEKASSVFNVNS	-	-	-	-	-	-	-
378-397	DIEKKIAKMEKASSVFNVN	++	++	++	++	++	++	++
378-396	DIEKKIAKMEKASSVFNV	++	++	++	++	++	++	++
378-395	DIEKKIAKMEKASSVFN	++	++	++	++	++	++	+
378-394	DIEKKIAKMEKASSVF	++	-	+	++	++	+	-
378-393	DIEKKIAKMEKASSV	++	-	-	++	++	-	-
378-392	DIEKKIAKMEKASS	++	-	-	++	+	-	-
378-391	DIEKKIAKMEKAS	-	-	-	++	-	-	-
378-390	DIEKKIAKMEKA	-	-	-	+	-	-	-
378-389	DIEKKIAKMEK	-	-	-	-	-	-	-

Peptides were synthesized, cleaved and purified by HPLC (see Table 3). T cells (2×10^4) of the clones shown in Table 1 were cultured with irradiated autologous EBV-B cells (10^4) in the presence of various antigen concentrations, ranging from 1 to 100 μ M. Data are presented as the concentration of peptide (μ M) required for 50% of maximum stimulation. ++, 50% of maximum stimulation obtained with 1 to 10 μ M of peptide; +, 50% of maximum stimulation obtained with 10 to 100 μ M peptide; -, any peptide that failed to stimulate proliferation at 100 μ M. Each clone is representative of a group of (at least four) CS.T3-specific clones with the same DR restriction. These results have been confirmed in three to four separate experiments.

Table 3 Antibody response in mice immunized with the B-cell epitope (NANP)₃ linked to the T-cell epitope CS.T3

Mouse strain	H-2	Anti-(NANP) ₅₀	Anti-sporozoite
C57BL/6	b	2,343	160
BALB/c	d	5,860	1,280
B10.MOla	f	938	40
C3H.HeJ	k	2,343	160
C3H.Q	q	3,705	320
B10.RIII	r	14,647	1,280
B10.SOla	s	1,482	1,280

Peptide (NANP)₃-CS.T3 was synthesized by the solid-phase technique using *N*-fluorenylmethoxy carbonyl (Fmoc)-amino acids, *t*-butyl-based side-chain protecting groups and a *p*-benzyloxybenzylalcohol polystyrene resin²⁴. The *N*^α-unprotected CS.T3 peptide resin was elongated by three repeated couplings of Fmoc-(NANP)₃ by the *O*-benzotriazolyl-*N,N,N,N*-tetramethyluronium hexafluorophosphate procedure²⁵ (Knorr *et al.* manuscript in preparation). The (NANP)₃-CS.T3 peptide was cleaved with trifluoroacetic acid/methylene chloride/anisole and purified by HPLC. Mice (two per group) were immunized at the base of the tail with 50 µg of (NANP)₃-CS.T3 in incomplete Freund's adjuvant. Eight weeks later, they were boosted with 25 µg of the immunogen in complete Freund's adjuvant. Plasma were taken between two and six weeks later and tested individually by enzyme-linked immunosorbent assay for the presence of anti-(NANP)₅₀ antibody, and by indirect immunofluorescence for antibodies to sporozoites^{26,27}. Enzyme-linked immunosorbent assay titres are geometric means of the last dilution of plasma with A₄₅₅ > 0.1 and > twice A₄₅₅ of plasma from mice injected with saline. The antigen used to coat the enzyme-linked immunosorbent assay plates was (NANP)₅₀. All pre-immune titres to (NANP)₅₀ were <150 and to sporozoites <40.

index²⁰; however, when the peptide is modelled as a regular α -helix, clusters of hydrophobic residues in the N- and C-terminal halves of the peptide are found on opposite sides of the helix (K. Muller, unpublished results). It will require more detailed information on the three-dimensional structure of MHC antigens²¹ to clarify the nature of the peptide binding.

Clones recognizing CS.T3 in association with DR2 and DR9 also recognize the parasite-derived CS protein^{5,22} as well as the cysteine-substituted CS.T3 peptide, whose sequence corresponds exactly to that of the CS protein. The same situation holds for at least four other DR antigens (DR1, 5, w6 and 7; ref. 28 and F. Sinigaglia *et al.*, unpublished results). Thus, this determinant might be an excellent candidate for a vaccine inducing cellular immunity to sporozoites. In addition to inducing potentially protective T-cell effector functions, the CS.T3 peptide can be used as a carrier for (NANP)_n. A vaccine consisting of the invariant CS.T3 T-cell epitope linked to an appropriate B-cell epitope might well induce both humoral and cellular parasite-specific immunity in the genetically diverse human population.

Received 11 July; accepted 1 November 1988.

- Benacerraf, B. & Germain, R. N. *Immunol. Rev.* **38**, 70-119 (1978).
- Buus, S., Sette, A., Colon, S., Miles, C. & Grey, H. M. *Science* **235**, 1353-1358 (1987).
- Del Portillo, H. A., Nussenzweig, R. S. & Enea, V. *Molec. Biochem. Parasitol.* **24**, 289-294 (1987).
- De la Cruz, V. F., Lal, A. A. & McCutchan, T. F. *J. biol. Chem.* **262**, 11935-11939 (1987).
- Sinigaglia, F. *et al. Eur. J. Immunol.* **18**, 633-636 (1988).
- Trucco, M. M., Garotta, G., Stocker, J. W. & Ceppellini, R. *Immunol. Rev.* **47**, 219-252 (1979).
- Watson, A. M., DeMars, R., Trowbridge, I. S. & Bach, F. H. *Nature* **304**, 358-361 (1983).
- Ziegler, A. & Milstein, C. *Nature* **279**, 243-244 (1979).
- Schwartz, R. H. A. *Rev. Immunol.* **3**, 237-261 (1986).
- Shastri, N., Gammon, G., Miller, A. & Sercarz, E. E. *J. exp. Med.* **164**, 882-896 (1986).
- Livingstone, A. M. & Fathman, C. G. A. *Rev. Immunol.* **5**, 477-502 (1987).
- Zavala, F., Cochrane, A. H., Nardin, E. H., Nussenzweig, R. S. & Nussenzweig, V. *J. exp. Med.* **157**, 1947-1957 (1983).
- Dame, J. B. *et al. Science* **225**, 593-599 (1984).
- Enea, V. *et al. Science* **225**, 628-630 (1984).
- Togna, R. *et al. J. Immun.* **137**, 2956-2960 (1986).
- Good, M. *et al. J. exp. Med.* **164**, 655-660 (1986).
- Del Giudice *et al. J. Immun.* **137**, 2952-2955 (1986).
- Baur, M. P. & Danilovs, J. A. in *Histocompatibility Testing* (ed. Terasaki, P.) 955-968 (UCLA, Los Angeles, 1980).

- Rothbard, J. & Taylor, W. *EMBO J.* **7**, 93-100 (1988).
- DeLisi, C. & Berzofsky, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7048-7052 (1985).
- Bjorkman, P. J. *et al. Nature* **329**, 506-518 (1987).
- Guttinger, M. *et al. EMBO J.* **7**, 2555-2557 (1988).
- Wilkinson, D. *et al. J. exp. Med.* **167**, 1442-1458 (1988).
- Atherton, A. & Sheppard, R. C. in *The Peptides: Analysis, Synthesis, Biology* Vol. 9. (eds Udenfriend, S. & Meienhofer, J.) 1-38 (Academic, New York, 1987).
- Dourtoglou, V., Ziegler, J. C. & Gross, B. *Tetrahedron Lett.* 1269-1272 (1987).
- Eitlinger, H. M. *et al. J. Immun.* **140**, 626-633 (1988).
- Eitlinger, H. M., Heimer, E. P., Trzeciak, A., Felix, A. M. & Gillesen, D. *Immunology* **64**, 551-558 (1988).
- Kilgus, J., Romagnoli, P., Guttinger, M., Adorini, L. & Sinigaglia, F. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Sequences homologous to *ZFY*, a candidate human sex-determining gene, are autosomal in marsupials

Andrew H. Sinclair*, Jamie W. Foster, James A. Spencer, David C. Page†, Mark Palmer‡, Peter N. Goodfellow‡ & Jennifer A. Marshall Graves

Department of Genetics and Human Variation, La Trobe University, Bundoora, Victoria 3083, Australia

† Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

‡ Human Molecular Genetics Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

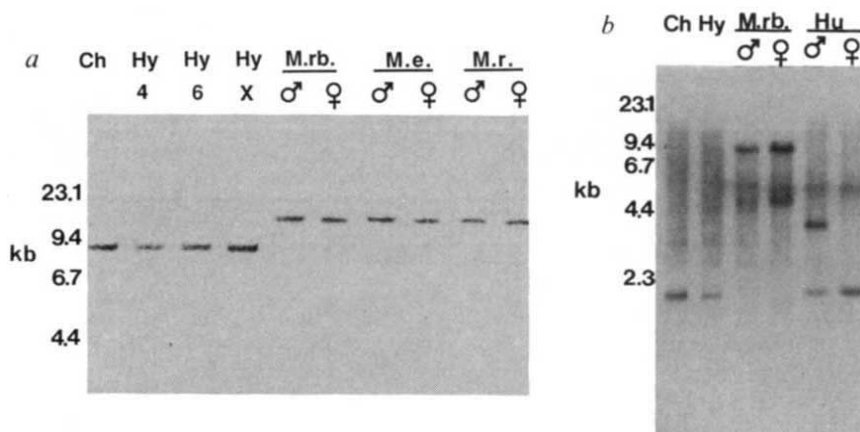
Sexual differentiation in placental mammals results from the action of a testis-determining gene encoded by the Y chromosome. This gene causes the indifferent gonad to develop as a testis, thereby initiating a hormonal cascade which produces a male phenotype^{1,2}. Recently, a candidate for the testis-determining gene (*ZFY*, Y-borne zinc-finger protein) has been cloned^{3,4}. The *ZFY* probe detects a male-specific (Y-linked) sequence in DNA from a range of eutherian mammals, as well as an X-linked sequence (*ZFX*) which maps to the human X chromosome³. In marsupials it is also the Y chromosome that seems to determine the fate of the gonad, but not all sexual dimorphisms⁵. Using the *ZFY* probe we find, surprisingly, that the *ZFY* homologous sequences are not on either the X or the Y chromosome in marsupials, but map to the autosomes. This implies *ZFY* is not the primary sex-determining gene in marsupials. Either the genetic pathways of sex determination in marsupials and eutherians differ, or they are identical and *ZFY* is not the primary signal in human sex determination.

Marsupials (infraclass Metatheria) diverged from placental mammals (infraclass Eutheria) at least 130 million years ago⁶, so comparisons of gene arrangements between these groups may provide information about the evolution of sex chromosomes and sex determination. Within the marsupials, the two orders Diprotodontia (represented in this work by *Macropus eugenii*, the tammar wallaby, and *M. robustus*, the wallaroo) and the Polyprotodontia (*Dasykaluta rosamondae* and *Sminthopsis crassicaudata*) diverged about 45 million years ago⁷.

In both these orders the marsupial X chromosome bears the genes *HPRT*, *G6PD*, *PGK* and *GLA*⁸ (symbols as in Fig. 3 legend), which are also X-linked in all eutherian mammals tested, confirming that at least part of the mammalian X is highly conserved in evolution, as predicted by Ohno¹. *STS* has been excluded from the marsupial X (ref. 9), however, and the recent localization of *DMD* (refs 10, 11) to one autosome (5p in *M. eugenii* and 3q in *D. rosamondae*) and *OTC* (ref. 12) and *SYN1* (refs 10, 11) to another (1p in *M. eugenii* and 3p in *D. rosamondae*) (symbols as in Fig. 3 legend), demonstrates that the region containing these genes is not present on the marsupial X. In humans these genes map to Xp21.2, Xp21.1 and Xp11.2,

* Present address: Human Molecular Genetics Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

Fig. 1 (a), Southern blot^{22,23} of *Pst*I-digested DNA samples probed with the human *ZFY* probe (pDP1007) and washed with $2\times$ SSC at 60 °C. DNA from male and female *M. robustus* (M.rb.), *M. eugenii* (M.e.), and *M. rufus* (M.r.) show a single band at ~12 kb, which is absent from cell hybrids (Hy) containing the *M. robustus* chromosomes 4, 6 and X on a Chinese hamster background. The bands at ~8 kb are the Chinese hamster (Ch) controls. b, Southern blot of *Eco*RI-digested DNA samples tested with human *ZFY* probe (pDP1007) and washed at low stringency ($3\times$ SSC at 50 °C). After overnight exposure at this stringency, one strongly hybridizing (~9.8 kb) band and a second weaker (~4.8 kb) band are revealed in DNA from *M. robustus* (M.rb.); a third (~1.8 kb) band is evident after longer exposure. All three M.rb. bands are present in both sexes and absent in DNA from a hamster-M.rb. cell hybrid (Hy) carrying M.rb. chromosomes 4, 6 and X. Ch is the hamster control, and Hu DNA is from a human male and female, showing a strong male-specific band, and a second band whose intensity in female is double that of male DNA.



respectively¹³, close to *ZFX*. Marsupial sex chromosomes are smaller than those of eutherians and apparently lack a pairing region^{14,15}. One possible explanation is that the smaller X and Y chromosomes of marsupials may have been derived by translocation of a large pseudoautosomal region from the partially differentiated X and Y chromosomes of a common ancestor to marsupial autosomes^{10,16}.

The probe pDP1007 (ref. 3) was hybridized to 'Noah's Ark' blots, containing DNA from male and female pairs of several species. The pDP1007 insert is a genomic sequence derived from the human Y chromosome, and encodes a putative 'zinc-finger' domain of *ZFY*. A single strong band was detected at high stringency, in a position identical for male and female, and identical across three *Macropus* species (Fig. 1a). Although there are clearly conserved marsupial sequences which show strong homology with human *ZFY*, these sequences are not male-specific (Y-linked) in any of these marsupial species. In addition, there was no difference in band intensity between male and female, as would be expected for X-linked sequences, suggesting that the sequences are autosomal. Single bands corresponding to different relative molecular masses were obtained with DNA from species belonging to other marsupial orders: *S. crassicaudata* and the American opossum, *Didelphis virginiana*. The absence of a *ZFX* homologue on the marsupial X was confirmed by probing Southern blots of DNA from hamster-marsupial cell hybrids which retained an intact *M. robustus* X chromosome⁹. DNA from the hybrid showed the band of ~8 kilobases (kb) detected in *Pst*I-digested hamster DNA controls, but not the band of ~12 kb detected in *M. robustus* DNA (Fig. 1a), confirming that a *ZFY* homologous sequence is not detected on the marsupial X, at least in this species. On Southern blots of *Eco*RI-digested marsupial DNA washed at low stringency, three *ZFY*-homologous bands were detected: one very strong, one weaker (Fig. 1b) and a third that could be detected only after prolonged exposure. Once again, these bands were of similar intensity and corresponded to the same molecular mass for male and female, and were absent from the cell hybrid.

These sequences were localized to *M. eugenii* and *S. crassicaudata* chromosomes by *in situ* hybridization with human *ZFY* probes pDP1007 (ref. 3) and pMF1. The plasmid pMF1 is a human testicular complementary DNA clone isolated using an oligomer derived from the published sequence of pDP1007; the 2.4-kb insert in pMF1 encodes the zinc-finger domain of pDP1007 and other coding sequences 5' to the finger region. These extra sequences contain an apparently complete open reading frame encoding an acid-rich protein domain. In Southern blot hybridizations of human DNA, pMF1 recognizes the same sequences as pDP1007 and additional Y-derived bands.

Analysis of the grain distribution revealed no hybridization site on either the X or Y in male or female of either species with either probe. But a strong hybridization signal (presumably equivalent to the single band on Southern blots at high stringency) was detected with both probes on chromosome 5p of both male and female *M. eugenii* (Fig. 2a) and on chromosome 3q of male and female *S. crassicaudata* (Fig. 2b) at the sites corresponding to the positions of the *DMD* sequences in these species^{10,11} (Fig. 3a, b). A secondary signal was consistently

Table 1 Analysis of *ZFY* *in situ* hybridization data for *M. eugenii* male

Chromosome	Deviance	Standardized residual
1	64	2.8
2	65	-2.2
3	63	-2.5
4	63	-2.4
5	29*	8.0*
6	66	-1.8
7	65	2.1
X	62	-2.4
Y	65	-1.8

A full account of this method of analysis is given elsewhere^{12,17}. The GLIM computer program was written¹⁷ to analyse the *in situ* hybridization data by comparing the observed distribution of grains with the distribution expected, omitting each chromosome in turn. The lowered deviance when the hybridized site is omitted provides a clear indication of the chromosome to which the gene has hybridized. An example of the output obtained in the present study (for *M. eugenii* male) is given in Table 1. Taking the data as a whole, the fit is adequate only for the model omitting chromosome 5 and the standardized residual is highly significant compared with the standard normal distribution. But from examination of the raw data we suspected that there could also be two weaker sites. To test whether these sites would hybridize significantly, we deleted chromosome 5 from the analysis and repeated the procedure with the remaining chromosomes. In the absence of chromosome 5, chromosome 1 has a deviance of 14 and a standardized residual of 5, and hence is a site of significant hybridization. Similarly, when both chromosomes 1 and 5 are deleted, then chromosome 7 shows a significant hybridization signal (deviance 4, standardized residual 5). Deletion of all three significant chromosomes from the analysis does not reveal any other significant sites of hybridization. Such analysis confirms chromosome 5 as the main site of hybridization, and chromosomes 1 and 7 as minor but significant sites in *M. eugenii*. Identical results were obtained for *M. eugenii* female cells. Similar analysis of the grain numbers of *S. crassicaudata* chromosomes confirm that 3q is the main site of hybridization and that 3p and 1p significant but minor sites in this species.

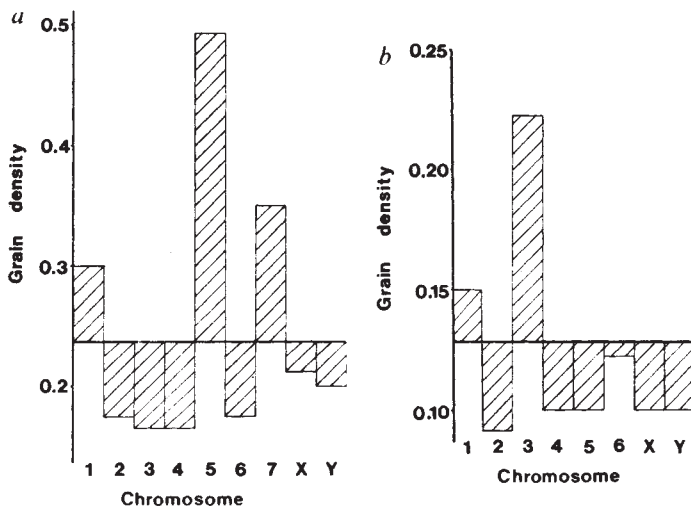


Fig. 2 Histograms showing results of *in situ* hybridization experiments. The line representing the expected number of grains per chromosome is calculated by dividing the total number of grains present over chromosomes by the relative proportion of the genome represented by each chromosome. *a*, Hybridization of pMF1 ($0.05 \mu\text{g ml}^{-1}$) to metaphase chromosomes from *M. eugenii* male. *b*, Hybridization of pDP1007 to metaphase chromosomes from *S. crassicaudata* male. The *in situ* hybridization method used for mapping heterologous probes has been reported elsewhere¹². Two independent experiments were carried out on each species, cells of both sexes were used, and slides were scored by two independent investigators, identifying chromosomes by size and morphology.

demonstrated on the *M. eugenii* chromosome 1p and the *S. crassicaudata* 3p, the location of the *OTC* and *SYN1* genes in these species¹⁰⁻¹². A third, weaker signal was consistently detected on *M. eugenii* chromosome 7 and *S. crassicaudata* 1p. Statistical analysis of the data with the GLIM program¹⁷ confirmed that these chromosomes contained one major and two minor sites of hybridization (Table 1). Proof of correlation between *in situ* hybridization data and Southern bands awaits cloning and localization of each sequence.

We conclude that *ZFY*-homologous sequences are autosomal in marsupials. The *ZFY* probe in marsupials maps to the same location as *DMD*, with a secondary site near the *OTC* and *SYN1* genes. All four of these genes are located close to *ZFX* on the human X chromosome. There is no evidence for either X or Y borne *ZFY*-homologous sequences in marsupials. Thus, none of the sequences can be the primary testis-determining gene, because this must logically be on the Y. Furthermore, the exclusion of these sequences from the marsupial X chromosome excludes the possibility that differences in their dosage between males and females can control other sexual dimorphisms.

The differences in the location of *ZFY* homologous sequences in eutherian and metatherian mammals must mean either that the primary sex-determination signal is accomplished by unrelated genes in the two mammalian infraclasses, or that *ZFY* is not the primary sex-determining gene in eutherians. The latter hypothesis is hard to reconcile with genetic and molecular analysis of the sex-determining region of the human Y.

In marsupials, as in eutherian mammals, the Y chromosome appears to be testis-determining; the few XXY animals described possess testes, whereas XO animals lack testes⁵. But it is clear that other sexual dimorphisms are a function of X-linked and/or autosomal genes. The phenotypes of marsupials with aneuploid sex chromosomes, unlike human XXY and XO, are not unambiguously male or female, suggesting that, whereas the Y chromosome determines testicular differentiation, the dosage of X chromosomes may influence the differentiation of scrotum, pouch and mammary glands⁵. Recent observations of early development show that the determination of scrotum, pouch, mammary glands, gubernaculum and processus vaginalis precedes gonadal differentiation¹⁸ and is not androgen-dependent, unlike that of eutherian mammals¹⁹⁻²¹. The differences in sex determination between marsupials and eutherians may conceivably relate to a more dominant role of *ZFY* in eutherians.

If testis determination turns out to be accomplished by unrelated genes in the two therian subclasses, it will be of interest to determine which was the ancestral system. The observation of a single band on Southern blots of DNA from male and female chickens³ implies that *ZFY*-homologous sequences may be autosomal in birds. This supports the proposition that the testis-determining function of *ZFY* has evolved since the divergence of the eutherian lineage 130 million years ago. The same proposition is indicated by the conservation between eutherian and marsupials of the close proximity of *ZFY* homologous sequences to *DMD* and *OTC*, *SYN1*, suggesting that the

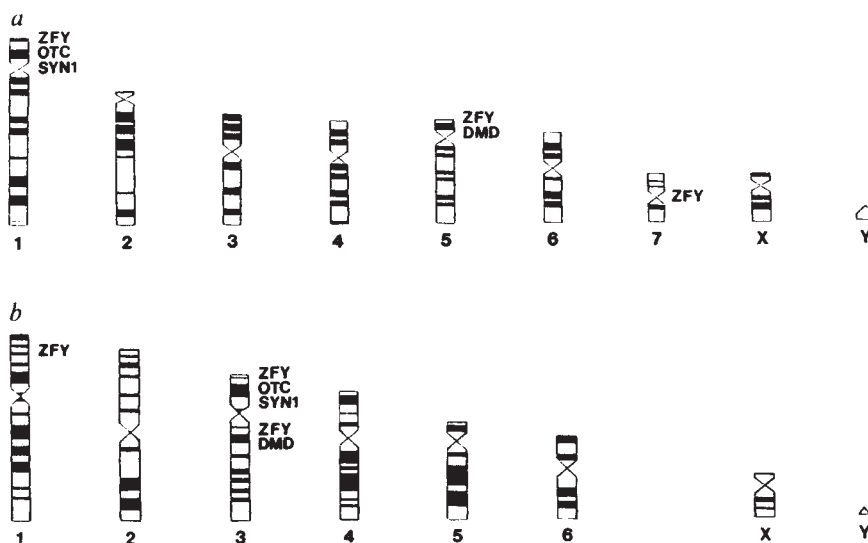


Fig. 3 Ideograms of *a*, *M. eugenii* and *b*, *S. crassicaudata* showing the sites of mapped genes. Gene symbols: *ZFY*, Y-linked zinc-finger protein; *ZFX*, X-linked zinc-finger protein; *DMD*, Duchenne muscular dystrophy; *OTC*, ornithine transcarbamylase; *SYN1*, synapsin I; *G6PD*, glucose-6-phosphate dehydrogenase; *HPRT*, hypoxanthine phosphoribosyl transferase; *PGK*, phosphoglycerate kinase; *GLA*, galactosidase- α ; *STS*, steroid sulphatase.

ancestral zinc-finger protein sequences must have been located near these genes in a common therian ancestor in a region that was pseudoautosomal or autosomal, and not subject to X-inactivation^{10,16}. If *ZFY*-homologous genes have a secondary role in sex determination in vertebrates other than eutherians, then a secondary step in a sex-determining pathway may have taken on a primary function in eutherians. The identification and isolation of the primary testis-determining gene, which must be borne on the Y in marsupials, now assumes great importance.

We thank Drs L. M. Kunkel and L. De Gennaro for use of their *DMD* and *SYN1* probes, respectively. This work was supported by an Australian Research Grants Scheme to J.M.G. and a National Institute of Health grant to D.L.P. We also thank Ms L. Gigliotti for typing the manuscript, and the Reprography Department, La Trobe University, for photography.

Received 27 October; accepted 15 November 1988.

- Ohno, S. *Sex Chromosomes and Sex-linked Genes* (Springer, Berlin, 1967).
- Jost, A., Vigier, B., Prepin, J. & Perchellet, J. P. *Recent Prog. Horm. Res.* **29**, 1-35 (1973).
- Page, D. C. *et al. Cell* **51**, 1091-1104 (1987).
- Page, D. C. *Phil. Trans. R. Soc. B* (in the press).
- Sharman, G. B., Robinson, E. S., Walton, S. M. & Berger, P. J. *J. reprod. Fert.* **21**, 57-68 (1970).
- Air, G. M., Thompson, E. O. P., Richardson, B. J. & Sharman, G. B. *Nature* **229**, 391-394 (1971).
- Archer, M. *Vertebrate Zoogeography and Evolution in Australasia* (eds Archer, M. & Clayton, G.) 780-781 (Hesperian, Western Australia, 1984).
- Graves, J. A. M. *Genetic Maps 4* (ed. O'Brien, S. J.) 504-507 (Cold Spring Harbor, New York, 1987).
- Dawson, G. W. & Graves, J. A. M. *Cytogenet. Cell Genet.* **42**, 80-84 (1986).
- Sinclair, A. H. thesis, La Trobe Univ., Australia (1988).
- Graves, J. A. M., Sinclair, A. H. & Spencer, J. A. *Aust. J. Zool.* (in the press).
- Sinclair, A. H., Wrigley, J. M. & Graves, J. A. M. *Genet. Res.* **50**, 131-136 (1987).
- Davies, K. E., Mandel, J. L., Weissenbach, J. & Fellous, M. *Cytogenet. Cell Genet.* **46**, 277-315 (1987).
- Hayman, D. L., Ashworth, L. K. & Carrano, A. V. *Cell Genet.* **34**, 365-370 (1982).
- Sharp, P. *Chromosoma* **86**, 27-47 (1982).
- Graves, J. A. M. *Trends Genet.* **3**, 252-256 (1987).
- Scott, D. J. & Sinclair, A. H. *Biometrics* (in the press).
- O. W.-S., Short, R. V., Renfree, M. B. & Shaw, G. *Nature* **331**, 716-717 (1988).
- Renfree, M. B. & Short, R. V. *Phil. Trans. R. Soc. B* (in the press).
- Wilson, J. D., George, F. W. & Griffin, J. E. *Science* **211**, 1278-1284 (1981).
- Kratochwil, K. *Dev. Biol.* **61**, 358-365 (1977).
- Southern, E. J. *J. molec. Biol.* **98**, 503-517 (1975).
- Reed, K. C. & Mann, D. *Nucleic Acids Res.* **13**, 7207-7221 (1985).

Cloning and expression of a rat D₂ dopamine receptor cDNA

James R. Bunzow, Hubert H. M. Van Tol,
David K. Grandy, Paul Albert, John Salon,
MacDonald Christie, Curtis A. Machida*,
Kim A. Neve† & Olivier Civelli‡

Vollum Institute for Advanced Biomedical Research,
L474, Oregon Health Sciences University,
3181 SW Sam Jackson Park Road, Portland, Oregon 97201, USA

* Division of Neuroscience, Oregon Regional Primate Center,
505 NW 185th Avenue, Beaverton, Oregon 97006, USA

† Veterans Administration Medical Center, 3710 SW
US Veterans Hospital Road, Portland, Oregon 97201, USA

Dopamine receptors are classified into D₁ and D₂ subtypes on the basis of their pharmacological and biochemical characteristics^{1,2}. The D₂ dopamine receptor has been implicated in the pathophysiology and treatment of movement disorders³, schizophrenia⁴ and drug addiction⁵. The D₂ dopamine receptor interacts with guanine nucleotide-binding proteins to induce second messenger systems^{6,7}. Other members of the family of receptors that are coupled to G proteins share a significant similarity in primary amino-acid sequence and exhibit an archetypical topology predicted to consist of seven putative transmembrane domains⁸. We have taken advantage of the expected nucleotide sequence similarities among mem-

Table 1 K_i values (nM) for L-RGB2Zem-1 and rat striatum

	RGB-2 transformed Ltk ⁻ cells	Rat striatum
(+)Butaclamol	0.83	1.0
(-)Butaclamol	> 1,000	> 1,000
Haloperidol	3.0	5.3
Dopamine + GTP	17,000	6,300
Sulpiride		
high affinity	80	67 (87%)
low affinity	—	> 10,000 (13%)
SCH 23390		
high affinity	—	35 (16%)
low affinity	1,000	780 (84%)
Ketanserin		
high affinity	—	27 (25%)
low affinity	> 1,000	1,000 (75%)

The 50% inhibitory concentration values (IC₅₀) calculated in Fig. 3b were converted to K_i values as described²⁴. Results are geometric means of three experiments in which 0.5 nM [³H]spiperone was inhibited by various concentrations of unlabelled drug. This concentration of radioligand is more than 10 times its K_d value. This explains the discrepancies between the IC₅₀ values in Fig. 3b and these K_i values. All binding data for L-RGB2Zem-1 membranes were fit best by assuming the presence of only one class of binding sites. On the other hand inhibition by several drugs of [³H]spiperone binding to rat striatal membranes was fit best by assuming the presence of two classes of binding sites. The proportions of binding sites with high and low affinity for inhibitors are shown in parentheses. SCH 23390 and ketanserin inhibited 10–20% of [³H]spiperone binding to rat striatal membranes with high affinity. In rat striatal membranes inhibition of radioligand binding by sulpiride was fit best by assuming one class of binding sites but 10–15% of the (+)butaclamol-displaceable binding was not inhibited by sulpiride at the concentrations used. It seems likely that the binding sites with high affinity for ketanserin and SCH 23390 and which are not displaced by sulpiride represent binding of [³H]spiperone to 5-HT₂ serotonin receptors in rat striatal membranes. Binding of SCH 23390 to 5-HT₂ receptors has been described previously²⁵. In rat striatal membranes the apparent affinity of drugs for D₂ dopamine receptors, the class of binding sites comprising 80–90% of [³H]spiperone binding, was indistinguishable from the apparent affinity of drugs to membranes prepared from L-RGB2Zem-1 cells. K_i values for the class of binding sites representing 10–25% of specific binding were calculated by assuming that the radioligand was binding to serotonin receptors with a K_d value of 1 nM.

bers of this gene family to isolate genes coding for new receptors. Using the hamster β₂-adrenergic receptor gene as a hybridization probe we have isolated related genes including a cDNA encoding the rat D₂ dopamine receptor. This receptor has been characterized on the basis of three criteria: the deduced amino-acid sequence which reveals that it is a member of the family of G-protein-coupled receptors; the tissue distribution of the mRNA which parallels that of the D₂ dopamine receptor; and the pharmacological profile of mouse fibroblast cells transfected with the cDNA.

A rat genomic library was screened under low-stringency hybridization conditions with a nick-translated 1.3-kilobase (kb) *Hind*III fragment containing most of the coding region of the hamster β₂-adrenergic receptor (β₂AR) gene⁹ (Fig. 1). Several clones were found to hybridize to the hamster probe. One clone, called RGB-2, contained a 0.8 kb *Eco*RI-*Pst*I fragment that hybridized to the hamster β₂AR probe in Southern blot analysis. This fragment was sequenced and shown to have a stretch of nucleotides with a high degree of identity (32 out of 40 bases) to the nucleotide sequence of transmembrane domain VII of the hamster β₂AR. One reading frame displayed a significant similarity to the amino-acid sequence of transmembrane domains VI and VII of the hamster β₂AR. This genomic fragment also contains a 3' intron splice site¹⁰ and 400 base pairs (bp) of putative intronic sequence (data not shown).

‡ To whom correspondence should be addressed.

The 0.8 kb *EcoRI*-*PstI* fragment was used to probe a rat brain cDNA library. Under high stringency hybridization conditions, two positive clones of about 2.5 kb were identified. Sequence and restriction analysis indicated that the two clones were replicas of a single clone and contained the sequence of the genomic fragment that had been used as a probe. When the RGB-2 cDNA was used as a hybridization probe in Northern blot analysis of mRNA isolated from rat brain, a band of approximately 2.5 kb was detected. This result indicated that the RGB-2 clone was nearly full length.

Figure 1a shows the nucleotide sequence of 2,455 bases for the RGB-2 cDNA. The longest open reading frame in this cDNA codes for a 415 amino-acid protein (relative molecular mass $M_r = 47,064$). This relative molecular mass is similar to that reported for the deglycosylated form of the D_2 dopamine receptor as determined by SDS-PAGE¹¹. An in-frame dipeptide that is 36 bases upstream from the putative initiation site is found in the 5' untranslated region of the RGB-2 cDNA. A small open reading frame has also been observed in the 5' untranslated sequence of the β_2 AR mRNA⁹.

Several structural features of the protein deduced from the RGB-2 cDNA demonstrate that it belongs to the family of G-protein-coupled receptors. First, a hydrophobicity plot of the protein sequence shows the existence of seven stretches of hydrophobic amino acids which could represent seven transmembrane domains⁸ (data not shown). Second, the primary amino-acid sequence of RGB-2 shows a high degree of similarity with other G-protein-coupled receptors (Fig. 1b). The regions of greatest amino-acid identity are clustered within the putative transmembrane domains. Within these domains the RGB-2 protein has a sequence identity of 48% with the human α_2 -adrenergic receptor¹², 41% with the human G-21 receptor¹³, 39% with the hamster β_2 -adrenergic receptor⁹, 27% with the porcine M_1 receptor¹⁴, and 26% with the bovine substance-K receptor¹⁵. Third, RGB-2 has several structural characteristics common to the members of the family of G-protein-coupled receptors. There are three consensus sequences for N-linked glycosylation in the N terminus with no signal sequence. Aspartate 80 found in transmembrane domain II is conserved in all known G-protein-coupled receptors. In transmembrane domain III Asp 114 corresponds to an Asp residue found in receptors that bind cationic amines¹⁶. Phosphorylation has been proposed as a means of regulating receptor function¹⁷. A potential site for phosphorylation by protein kinase A exists at Ser 228 in the third cytoplasmic loop. In the C-terminal portion of the rhodopsins and β -adrenergic receptors, there are many Ser and Thr residues which are potential substrates for receptor kinases¹⁸; in contrast, the short C terminus of RGB-2 has no Ser and Thr residues. As for the α_2 -adrenergic receptor however, RGB-2 has many Ser and Thr residues (22) in the third cytoplasmic loop which could serve as phosphorylation substrates. RGB-2 contains a large cytoplasmic loop (135 amino acids) between transmembrane domains V and VI with a short C terminus (14 amino acids). This structural organization is similar to other receptors which are coupled by G_i such as the α_2 -adrenergic receptor and the M_2 muscarinic receptor. Unlike the members of the adrenergic and muscarinic receptor families, the RGB-2 gene has at least one intron in its coding sequence which is located in transmembrane domain VI.

As a first step towards determining the identity of RGB-2, the tissue distribution of the RGB-2 mRNA was examined by Northern blot analysis (Fig. 2). The RGB-2 mRNA is expressed at different levels in various regions of the rat brain with the basal ganglia showing the highest concentration. Furthermore, the RGB-2 mRNA was found in high amounts in the neurointermediate lobe of the pituitary gland of the rat and to a lesser degree in the anterior lobe of this gland. The expression pattern of the RGB-2 mRNA is strikingly similar to the distribution of the D_2 dopamine receptor as determined by receptor autoradiography and binding studies of tissue preparations¹⁹.

To study the pharmacological characteristics of the receptor encoded by RGB-2, the cDNA was expressed in eukaryotic cells. The RGB-2 cDNA was cloned into the eukaryotic expression vector pZem3 which initiates transcription from the mouse metallothionein promoter²⁰. This plasmid was co-transfected with the selectable neomycin phosphotransferase gene (pRSVneo) into the Ltk⁻ mouse fibroblast cell line and stable transfectants were isolated. Expression of the RGB-2 mRNA was verified by Northern blot analysis and a cell line expressing RGB-2 designated L-RGB2Zem-1 was isolated. The RGB-2 mRNA was not detectable in the parent Ltk⁻ cell line.

As the RGB-2 mRNA displayed the tissue distribution expected of the D_2 dopamine receptor, we performed a pharmacological study of the L-RGB2Zem-1 cell line, native Ltk⁻ cells and rat striatum using the D_2 ligand [³H]spiperone. Membranes prepared from control Ltk⁻ cells showed no (+)butaclamol- or sulpiride-displaceable binding of [³H]spiperone. Binding of [³H]spiperone to membranes prepared from L-RGB2Zem-1 cells was saturable with a dissociation constant (K_d) of 48 pM (Fig. 3a). This value agrees with that observed for binding of [³H]spiperone to rat striatal membranes in parallel experiments (52 pM). The density of binding sites as determined in four

Fig. 1 Sequence of the rat D_2 dopamine receptor and comparison with the sequences of other G protein-coupled receptors. *a*, Nucleotide sequence for the RGB-2 cDNA and deduced amino-acid sequence of the rat D_2 dopamine receptor. The nucleotide sequence is numbered beginning with the first methionine of the long open reading frame. Nucleotide numbering is beneath the nucleotide sequence at the right hand end of each line. The deduced amino-acid sequence is shown above the nucleotide sequence. The double underline denotes the small open reading frame in the 5' untranslated region. Postulated N-linked glycosylation sites are indicated by asterisks. Putative protein kinase A phosphorylation sites have a line above them. The intron splice junction is designated by an arrow. The poly(A) adenylation site is underlined. *b*, Alignment of the amino-acid sequences of the rat D_2 dopamine receptor, the hamster β_2 -adrenergic receptor, the human α_2 -adrenergic receptor, the human G-21 receptor, the porcine muscarinic M_1 receptor and the bovine substance-K receptor (SK). Amino acids enclosed in solid lines and shaded represent residues that are identical in at least two other sequences compared to the D_2 dopamine receptor. The putative transmembrane domains are indicated by brackets labelled by roman numerals. The number of residues in the variable third cytoplasmic loop and at the C terminus are in parentheses.

Methods. The hamster β_2 AR gene was cloned from a partial hamster lung λ gt10 genomic DNA library (constructed from size-fractionated *EcoRI*-digested DNA) with two oligonucleotide probes designed from the published sequence⁹. The 1.3-kb *HindIII* fragment of the hamster β_2 AR gene which contains most of the coding sequence of that gene was labelled by nick-translation and used to probe a rat genomic DNA library in EMBL3. The library was transferred to Colony Plaque Screen filters (NEN) and screened with the ³²P-labelled probe under these hybridization conditions: 25% formamide, 5×SSC, 5×Denhardt's, 0.1% sodium pyrophosphate, 1% SDS and salmon sperm DNA (100 μ g ml⁻¹) at 37 °C. Filters were washed in 2×SSC and 0.1% SDS at 55 °C. One clone called RGB-2 was partially sequenced and found to be related to the hamster β_2 -adrenergic receptor. A nick-translated 0.8 kb *EcoRI*-*PstI* fragment from RGB-2 was used to screen a rat brain cDNA library in λ gt10 with the same hybridization conditions as above except that 50% formamide was used. The filters were washed in 0.2×SSC and 0.1% SDS at 65 °C. Two positive clones were isolated from a library of 500,000 clones. DNA sequence was obtained from both strands by the Sanger dideoxy chain determination method using Sequenase (US Biochemical Corporation)²⁶.

a

GGCTGGCG
-120
SAGGGGCGGCGTGGCTGGATGCGGGGAGCTGGAAGCTCGAGCAGCGGGCGCTTCTTGGCCCGGGCCCATATGGCTTCAAGAGCGGTGCCACCCAGTGGCCCACTGGCCCA
-1
1 10 20 30
Met Asp Pro Leu Asn Leu Ser Trp Tyr Asp Asp Asp Leu Glu Arg Gln Asn Trp Ser Arg Pro Phe Asn Gly Ser Glu Gly Lys Ala Asp
ATG GAT CCA CTG AAC CTG TCC TGG TAC GAT GAC GAT CTG GAG AGG CAG AAC TGG AGC CGG CCC TTC AAT GGG TCA GAA GGG AAG GCA GAC
1 90
Arg Pro His Tyr Asn Tyr Tyr Ala Met Leu Leu Thr Leu Leu Ile Phe Ile Ile Val Phe Gly Asn Val Leu Val Cys Met Ala Val Ser
AGG CCC CAC TAC AAC TAC TAT GCC ATG CTG CTC ACC CTC CTC ATC TTT ATC ATC GTC TTT GGC AAT GTG CTG GTG TGC ATG GCT GTA TCC
180
70 80 90
Arg Glu Lys Ala Leu Gln Thr Thr Thr Asn Tyr Leu Ile Val Ser Leu Ala Val Ala Asp Leu Leu Val Ala Thr Leu Val Met Pro Trp
CGA GAG AAG GCT TTG CAG ACC ACC ACC AAC TAC TTG ATA GTC AGG CTT GCT GTG GAT CTT CTG GTG GCC ACA CTG GTA ATG CCG TGG
270
100 110 120
Val Val Tyr Leu Glu Val Val Gly Glu Trp Lys Phe Ser Arg Ile His Cys Asp Ile Phe Val Thr Leu Asp Val Met Met Cys Thr Ala
GTT GTC TAC CTG GAG GTG GTG GGT GAG TGG AAA TTC AGC AGG ATT CAC TGT GAC ATC TTT GTC ACT CTG GAT GTC ATG ATG TGC ACA GCA
360
130 140 150
Ser Ile Leu Asn Leu Cys Ala Ile Ser Ile Asp Arg Tyr Thr Ala Val Ala Met Pro Met Leu Tyr Asn Thr Arg Tyr Ser Ser Lys Arg
AGC ATC CTG AAC CTG TGT GCC ATC AGC ATT GAC AGG TAC ACA GCT CTG GCA ATG CCC ATG CTG TAT AAC ACA CGC TAC AGC TCC AAG CGC
450
160 170 180
Arg Val Thr Val Met Ile Ala Ile Val Trp Val Leu Ser Phe Thr Ile Ser Cys Pro Leu Leu Phe Gly Leu Asn Asn Thr Asp Gln Asn
CGA GTT ACT GTC ATG ATT GCC ATT GTC TGG GTC CTG TCC TTC ACC ATC TCC TGC CCA CTG CTC TTC GGA CTC AAC AAT ACA GAC CAG AAT
540
190 200 210
Glu Cys Ile Ile Ala Asn Pro Ala Phe Val Val Tyr Ser Ser Ile Val Ser Phe Tyr Val Pro Phe Ile Val Thr Leu Leu Val Tyr Ile
GAG TGT ATC ATT GCC AAC CCT GCC TTT GTG GTC TAC TCC TCC ATT GTC TCA TTC TAC GTG CCC TTC ATC GTC ACT CTG CTG GTC TAT ATC
630
220 230 240
Lys Ile Tyr Ile Val Leu Arg Lys Arg Arg Lys Arg Val Asn Thr Lys Arg Ser Ser Arg Ala Phe Arg Ala Asn Leu Lys Thr Pro Leu
AAA ATC TAC ATC GTC CTC CGG AAG CGC CGG AAG CGG GTC AAC ACC AAG CGC AGC AGT CGA GCT TTC AGA GCA AAC CTG AAG ACA CCA CTC
720
250 260 270
Lys Asp Ala Ala Arg Arg Ala Gln Glu Leu Glu Met Glu Met Leu Ser Ser Thr Ser Pro Pro Glu Arg Thr Arg Tyr Ser Pro Ile Pro
AAG GAT GCT GCC CGC CGA GCT CAG GAG CTG GAA ATG GAG ATG CTG TCA AGC ACC AGC AGC CCC CCA GAG AGG ACC CGG TAT AGC CCC ATC CCT
810
280 290 300
Pro Ser His His Gln Leu Thr Leu Pro Asp Pro Ser His His Gly Leu His Ser Asn Pro Asp Ser Pro Ala Lys Pro Glu Lys Asn Gly
CCC AGT CAC CAC CAG CTC ACT CTC CCT GAT CCA TCC CAC CAC GGC CTA CAT AGC AAC CCT GAC AGT CCT GCC AAA CCA GAG AAG AAT GGG
900
310 320 330
His Ala Lys Ile Val Asn Pro Arg Ile Ala Lys Phe Phe Glu Ile Gln Thr Met Pro Asn Gly Lys Thr Arg Thr Ser Leu Thr Thr Met
CAC GCC AAG ATT GAT CCC AGG ATT GCC AAG TTT TTT GAG ATC CAG ACC ATG CCC AAT GGC AAA ACC CGG ACC TCC CTT AAG AGC ATG
990
340 350 360
Ser Arg Arg Lys Leu Ser Gln Gln Lys Glu Lys Lys Ala Thr Gln Met Leu Ala Ile Val Leu Gly Val Phe Ile Ile Cys Trp Leu Pro
AGC CGC AGA AAG CTC TCC CAG CAG AAG GAG AAG AAA GCC ACT CAG ATG CTT GCC ATT GTT CTC GGT GTG TTC ATC ATC TGC TGC CTG CCC
1080
370 380 390
Phe Phe Ile Thr His Ile Leu Asn Ile His Cys Asp Cys Asn Ile Pro Pro Val Leu Tyr Ser Ala Phe Thr Trp Leu Gly Tyr Val Asn
TTC TTC ATC ACG CAC ATC CTG AAT ATA CAC TGT GAT TGC AAC ATC CCA CCA GTC CTC TAC AGC GCC TTC ACA TGG CTG GGC TAT GTC AAC
1170
400 410 415
Ser Ala Val Asn Pro Ile Ile Tyr Thr Thr Phe Asn Ile Glu Phe Arg Lys Ala Phe Met Lys Ile Leu His Cys
AGT GCC GTC AAC CCC ATC ATC TAC ACC ACC TTC AAC ATC GAG TTC CGC AAG GCC TTC ATG AAG ATC TTG CAC TGC TGAGTCTGCCCTTGCCTG
1264
CACAGCAGCTGCTTCCACCTCCCTGCCATGACGGCCAGACCTCATCCCTGCAAGCTTGGGCGAAGAGGCCAGATGAACCTGGCTTCTCTCGACCTCGACGCCCTGCAGTGTGA
1383
GCTTGGCTGCATGCCCTCTCTGCCACACACCTCATCTGCCAGGTAGGCGCAGGAGACTGGTATCTTACACGCTTGGGCTTGGACCCATGGCTCAGGCGACCTCACAGATGTC
1502
CCCTCTCATATCCAGACCTCTCTCTGGACCAAGATGACGGCGCTTCTTTCGACCTTCTCTTGGGACAGAACTAGCTCAGTGGTGAGCACACCTGATCGCTGGCTTGGCC
1621
TGGCCCTTGGCTTGGCTGTGCCGATCAGGTGGTGGGAGGAGCAGCTTCTTACTTTATAGGAACACATAGGAAGCAGGGAACAGCCAGTCTCCAGGCAACATCAGTGTGAGG
1740
AGACACACATAAACACAGGTAGCTCCTGACCCAGAGAACTGAGGCTGAAAAATCTCTTTTCCACTCCAACTAGTGTGAGTCCCTACTTTTATAGCCATGGGTATTACTATG
1859
TCCTACTCTGTTATAGTATCCATGGGTTTCTGTACCATTTGGGGAAAACTCTTAATCTCAAGGGCCCCAAGAACTGTGAAGGAGAAAAATAGGCTGATCTCCCTTACTCT
1978
CCAATCCACTCCCACTCTCTGATATACCTTGATGTATCATCTCTCACAGCAATGCTGGCCAGTCAGGCGCTTGGACAGTGTGGAGTGAAGCTGGATGTGGTAACCTGGGGCT
2097
CTTTGGGCTGGGGGGTGTGAACATGCTCTCTCTTCCATATCTCTTCTTCCAGTCCCTTGGCTTAGAAGAGGCTGTGGTGGGGTGTGGAGTGTGATACCATTTGGGCTGG
2216
CCCTGAATGAGGAGGGAGCTGCACTTTGGAGGTTCTGGGATCCAACTCTGAACATCACTATACCTGTACCAAACTAATAAAACCTTGACAAGAGTCAAAAAA
2317

b

I II
D₂ MD...PL...NLSWYDDDLERQNSRPFNGSEKADRPYNYAMITLL IF IIVFGNVLVCMASREKALQTTINYLIVSLAVADIIVATLWMPVVYLEMGEKFSRIH
β₂ MCP...P...QNDSDFLTNGSHV...PDHDTVTEERDEAWVVGMAHMSVIVLAIVFGNVLVITAIKFERLOTTINYLIVSLAVADIIVATLWMPVVYLEMGEKFSRIH
α₂ MGSLO.PDA.GNASWNGTEAPG..GGARATPYSLOVT...LTLVLCLL.MLLTVEGNYLVIIIVATFSRAIKAPQNLFIIVSLASAHIIIVATLWIPFSLANEVMEWYCKTW
G-21 MDVLS.PGO.QNNTTSPAPPE..TGGNTTGISDVTVSQV.ITSLIGTL.IFCVTLGNACVVAALALERSLQNVANYLEGSLAVDIMUSLVIVLEMAALYQVINKWTLGQVT
M₁ MNTSAPPVSHNITVLAP.....GKGPPQVA.....FIGITTLG.LSLATVIGNLVIIISFKVNTLKTIVNNYFLISLACADIIIGTSMNLTYTYLLMHWALGTLA
SK MGACV.VMTDINTS.....SGLDSNATGITAFSMPGWQLALWTAAYLAL.VLVAVMGATVWIIILAHQMRITVNNYFEVNLALADICMAAFNAAFNFVYASHNINVEGRAF

III IV V
D₂ CDIFVTLDDVMCTASILNLCAISIDRYTAVAMPMLYNTTRYSSKRVTVMTAIVHVISLITSC.PLLFGLNNTD.....QNECIIANPAFVYSSIVSEFYVFIIVTLVYI
β₂ CEFWTSIDVLCVFASITELCVIADRYTATITSEFKYQSLLTKNKARMMLLMVMIVSGLTSLFIQIMHWYRATH..QKADICYHRETCDDFTNQAYATASSIVSFYVFIIVMV
α₂ CEIYLALDVLCTSSIVHLCASIDRYNSITQATYENLKRTPPRRIKAI.IITVWVISAIVISFPPLISIEKKGGG....GGPQPAEPCEINDQKWYVSSCIGSFAPQIMIM
G-21 COLFIALDVLCTSSILHLCASIDRYNITDPIIDYVINKRTPPRRALT.SLT.MLIGELISIPMLGWRTPEDR.....SDPDACTISKDMGYTIVSTFGATYIPELML
M₁ CDLWLALDYVASNASVMNLLISIDRYFSVTRPLSYRAKRTPRRAALM.LSLANLVSEVLWA.PAILFWQYLVGE...RTVLAGQCYIQFLSQPIITFGTAMAFYLPVIMVC
SK CYFQNLFPITAMFVSIYSMTATAADRYMAIVHVFQPRLSAPGTB..AV.LRGITLVLVALALAF.PQCFYSTITTTDEGATKCVVAPEDSGGKMLLLYHLIVIALIYE.LPVIMVF

VI VII
D₂ KIVIVLRKRKRVTNKR-(111)-KEKKATQMLALVLCMFILCWLPFFITTHLNIHDCN....IPVLSAFTWLCGYVNSA..VNPILTYTTFNFERNAFMKILHC
β₂ FVYSRVFQVAKRQIKI-(32)-KEHKALKTLGLIMGTFLCWLPFFIVNIVHVIDNLN....IPKEVILLNLWLCGYVNSA..FNPLTYCRSP.DERTAFQELL.CL-(36)
α₂ LVYVRIYQIAKRTRVP-(137)-REKRTFVLAIVLCMFVVCFFFTYTFVAVGCS....VETTLKFFFWFCYCNSS..INPMVITIFNHDERRAFKKILCRC-(6)
G-21 VLYGRIFRAARFRIPK-(110)-REKRTVKTGLIMGTFLCWLPFFIVNIVHVIDNLN....IPKEVILLNLWLCGYVNSL..INPMVITAYENKDFQNAFKKILKCNFCRQ
M₁ TLYNRIYRETNAREL-(137)-KEKKAARTLSAILLAFIVTMVFNLMVLSTFCKDC....VETTLWELGYWLCYVNST..INPMVIALCNKAFRTDTERLLILCR-(24)
SK VAYSVIGLTLVRRSVPG-(12)-ARKKFKVTMVLVIVTFAICWLYHYLYFLGTFTQEDYCHKFIQVYLALFWIA..MSTMYNIIITCCINHRFRSGERLAFRC-(62)

Fig. 2 Northern blot analysis of RGB-2 transcripts in rat brain and pituitary. Each lane contained 20 μ g of total RNA. Numbers on the right indicate kb as determined from RNA size markers (BRL). Lanes 1, olfactory bulb; 2, hippocampus; 3, cerebellum; 4, posterior cortex; 5, anterior cortex; 6, thalamus; 7, hypothalamus; 8, medulla; 9, amygdala; 10, mesencephalon; 11, septum; 12, posterior basal ganglia; 13, anterior basal ganglia; 14, neurointermediate lobe of pituitary; 15, anterior lobe of the pituitary.

Methods. Brain tissue was dissected from male Sprague Dawley rats and RNA isolated as described^{27,28}. For Northern blotting RNA was denatured using glyoxal and dimethyl sulphoxide and run on 1.2% agarose gels. After electrophoresis RNA was blotted onto a nylon membrane (N-bond, Amersham) and baked for 2 h at 80 °C. The membranes were pre-hybridized in 50% formamide, 0.2% polyvinyl pyrrolidone (M_r 40,000), 0.2% ficoll (M_r 400,000), 0.05 M Tris pH 7.5, 1 M NaCl, 0.1% PPI, 1% SDS and denatured salmon sperm DNA (100 μ g ml⁻¹) for 16 h at 42 °C. A random-primed ³²P labelled fragment (1×10^9 – 2×10^9 d.p.m. μ g⁻¹) from the 1.6 kb *Bam*HI–*Bgl*III fragment of RGB-2 which contains the coding region of this clone was used at 10^7 d.p.m. ml⁻¹ in the hybridization solution from above to probe the filters overnight at 42 °C. The blots were washed twice in $2 \times$ SSC and 0.1% SDS at 65 °C for 10 min, twice in $0.5 \times$ SSC and 0.1% SDS at room temperature for 15 min and once in $0.1 \times$ SSC and 0.1% SDS at 65 °C for 15 min. Blots were exposed overnight at 80 °C to X-ray film with an intensifying screen.

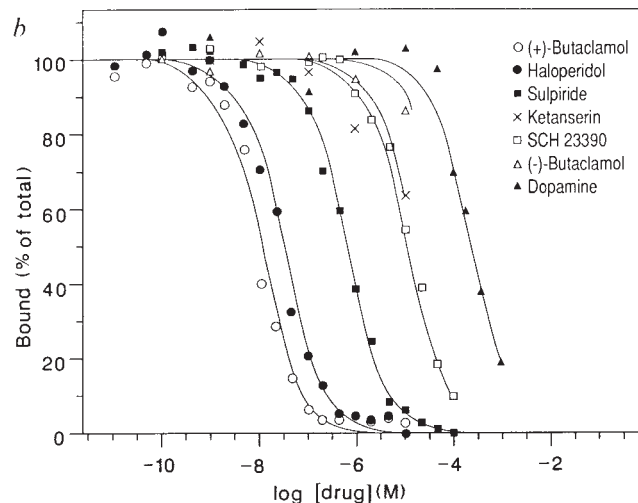
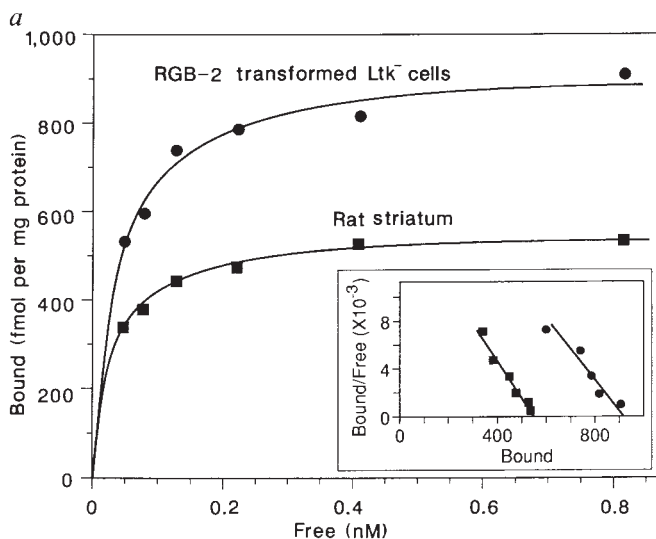
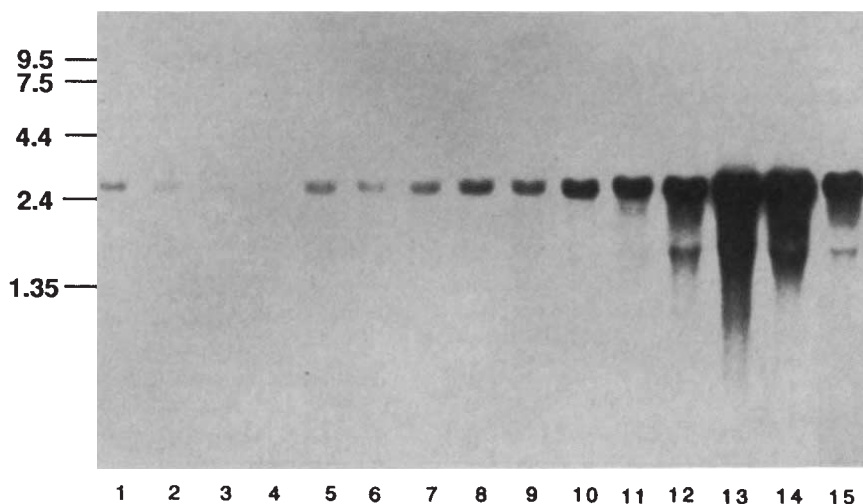


Fig. 3 Binding of [³H]spiperone to membranes from L-RGB2Zem-1 cells. *a*, Saturation isotherms of the specific binding of [³H]spiperone to membranes prepared from L-RGB2Zem-1 cells and rat striatum. Results are shown from one of four independent experiments. Inset: Scatchard transformation of the data. In this experiment K_d and B_{max} values for membranes prepared from L-RGB2Zem-1 were 40 pM and 876 fmol per mg protein, whereas the corresponding values in striatal membranes were 37 pM and 547 fmol per mg protein. *b*, Competition curves using L-RGB2Zem-1 membranes. Representative curves are shown for inhibition of specifically bound [³H]spiperone by drugs in membranes from L-RGB2Zem-1 cells. Each drug was tested three times.

Methods. The full RGB-2 cDNA was cloned into the plasmid pZem3²⁹. This plasmid was co-transfected with the plasmid pRSVneo into mouse Ltk⁻ cells by the standard CaPO₄ precipitation technique²⁰. Cells were selected in 350 μ g ml⁻¹ of G418. Transfectants were isolated and checked for expression of RGB-2 mRNA by Northern blot analysis. Membranes were prepared by homogenizing cells with a Dounce homogenizer at 4 °C in 0.25 M sucrose, 25 mM Tris pH 7.4, 6 mM MgCl₂, 1 mM EDTA. The homogenizing solution was centrifuged at 800g for 10 min and the pellet was subjected to a second homogenization and centrifugation. The supernatants were pooled and centrifuged at 200,000 g for 1 h. The pellet of this centrifugation was resuspended in 25 mM Tris pH 7.4, 6 mM MgCl₂, 1 mM EDTA at approximately 250 μ g protein ml⁻¹ and stored in small aliquots at -70 °C. Radioligand binding assays were carried out in duplicate in a volume of 2 ml (saturation analyses) or 1 ml (inhibition curves) containing (final concentration): 50 mM Tris, pH 7.4, 0.9% NaCl, 0.025% ascorbic acid, 0.001% bovine serum albumin, [³H]spiperone (Amersham, 95 Ci mmol⁻¹) and appropriate drugs. In some experiments 100 μ M guanosine 5'-triphosphate was included. (+)-Butaclamol (2 μ M) was used to define nonspecific binding. Incubations were initiated by the addition of 15–40 μ g of protein, carried out at 37 °C for 50 min, and stopped by the addition of 10 ml of ice-cold wash buffer (10 mM Tris, pH 7.4 and 0.9% NaCl) to each assay. The samples were filtered through glass-fibre filters (Schleicher & Schuell No. 30) and washed with an additional 10 ml of wash buffer. The radioactivity retained on the filter was counted using a Beckman LS 1701 scintillation counter. Data were analysed as previously described³⁰

except that curves were drawn using the data analysis program Enzfitter.

experiments was 945 fmol per mg of protein in L-RGB2Zem-1 membranes and 454 fmol per mg of protein in rat striatal membranes. The binding of [3 H]spiperone to membranes from L-RGB2Zem-1 cells was inhibited by a number of drugs and the resulting K_i values closely matches those obtained using striatal membranes (Fig. 3b, Table 1). The D_2 antagonists (+)butaclamol and haloperidol were the most potent inhibitors, followed by sulpiride. The D_1 dopamine antagonist SCH 23390 and the 5-HT $_2$ antagonist ketanserin were much less potent at blocking [3 H]spiperone binding. The binding seemed to be stereoselective as (+)butaclamol was much more potent than (-)butaclamol at inhibiting binding. In these experiments the absolute affinities of dopaminergic antagonists and the rank order of potency of the drugs (spiperone > (+)butaclamol > haloperidol > sulpiride > (-)butaclamol) agree closely with previously published values for the D_2 dopamine receptor².

The physiological effects of stimulation of D_2 dopamine receptors seem to be mediated by G_i . Inhibition of agonist binding to D_2 dopamine receptors by GTP is thought to be due to GTP-induced dissociation of a receptor- G_i complex which has a high affinity for agonist^{21,22}. Binding of [3 H]spiperone was inhibited by the agonist dopamine with a K_i of 17 μ M and a Hill coefficient of 1 in L-RGB2Zem-1 membrane. This dopamine binding was not responsive to the addition of GTP. This finding is consistent, however, with the reported lack of G_i in L cells²³. The pharmacological data presented here prove that the binding profile of the D_2 dopamine receptor is found in Ltk⁻ cells expressing the RGB-2 cDNA.

When transfected into eukaryotic cells, the RGB-2 cDNA directs the expression of a D_2 dopamine-binding protein. As the mRNA corresponding to this cDNA is localized in tissues where the D_2 dopamine receptor is known to be present and as this mRNA codes for a protein that has all the expected characteristics of a G-protein-coupled receptor, we conclude that RGB-2 is a clone for the rat D_2 dopamine receptor. The successful cloning of the D_2 dopamine receptor will provide new tools to study the regulation and function of this receptor.

We thank our colleagues of the VIABR for help and discussions. We are indebted to Alan North, Eckard Weber, Michael Forte, and Mike Lacey for manuscript review and Dee Yarozecki for manuscript preparation. D.K.G. and P.A. hold fellowships from the NIH and H.H.M.V.T. from the Canadian Heart Association. KAN has a grant from the Medical Research Foundation of Oregon. This work was supported by a NIH grant and a grant from Cambridge Neuroscience Research Inc., Cambridge, Massachusetts, to O.C.

Received 21 September; accepted 7 November 1988.

- Creese, I., Sibley, D. R., Hamblin, M. W. & Leff, S. E. *A. Rev. Neurosci.* **6**, 43-71 (1983).
- Seeman, P. *Pharmac. Rev.* **32**, 229-313 (1980).
- Lee, T. et al. *Nature* **273**, 59-61 (1984).
- Seeman, P. et al. *Science* **225**, 728-731 (1984).
- Barnes, D. M. *Science* **241**, 415-417 (1988).
- Cote, T. E., Frey, E. A., Grewe, C. W., & Keabian, J. W. *J. Neural. Trans. Suppl.* **18**, 139-147 (1983).
- Senogles, S. E. et al. *J. Biol. Chem.* **262**, 4860-4867 (1987).
- Dohlman, H. G., Caron, M. G. & Lefkowitz, R. J. *Biochemistry* **26**, 2657-2664 (1987).
- Dixon, R. A. F. et al. *Nature* **321**, 75-79 (1986).
- Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
- Grigoriadis, D. E., Niznik, H. B., Jarvie, K. R. & Seeman, P. *FEBS Lett.* **227**, 220-224 (1988).
- Kobilka, B. K. et al. *Science* **238**, 650-656 (1987).
- Kobilka, B. K. et al. *Nature* **329**, 75-79 (1987).
- Kubo, T. et al. *Nature* **323**, 411-416 (1986).
- Masu, Y. et al. *Nature* **329**, 836-838 (1987).
- Strader, C. D. et al. *J. Biol. Chem.* **263**, 10267-10271 (1988).
- Sibley, D. R., Benovic, J. L., Caron, M. G. & Lefkowitz, R. J. *Cell* **48**, 913-922 (1987).
- Bouvier, M. et al. *Nature* **333**, 370-373 (1988).
- Boyson, S. J., McGonigle, P. & Molinoff, P. B. *J. Neurosci.* **6**, 3177-3188 (1986).
- Gorman, C., Padmanabhan, R. & Howard, B. H. *Science* **221**, 551-553 (1983).
- Hamblin, M. W., Leff, S. E. & Creese, I. *Biochem. Pharmac.* **33**, 872-877 (1984).
- Dolphin, A. C. *Trends Neurosci.* **10**, 53-57 (1987).
- Jones, S. V. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4056-4060 (1988).
- Cheng, Y. C. & Prusoff, W. H. *Biochem. Pharmac.* **22**, 3099-3108 (1973).
- Hyttel, J. *Eur. J. Pharmac.* **91**, 153-154 (1983).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

- Chirgwin, J. M., Przybyla, A. E., McDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
- Ullrich, A. et al. *Science* **196**, 1313-1319 (1977).
- Uhler, M. & McKnight, G. S. *J. Biol. Chem.* **262**, 15202-15207 (1987).
- Neve, K. A. & Molinoff, P. B. *Molec. Pharmac.* **30**, 104-111 (1986).

A novel cancer therapy using a Mössbauer-isotope compound

Randell L. Mills*, Carl W. Walter†, Lata Venkataraman‡, Kevin Pang† & John J. Farrell§

* Mills Technologies Inc., RD No. 2, Cochranville, Pennsylvania 19330, USA

† Harvard Medical School, 10 Shattuck Street, Boston, Massachusetts 02115, USA

‡ Brigham and Women's Hospital, Harvard Medical School, Boston Massachusetts 02115, USA

§ Chemistry Department, Franklin and Marshall College, Lancaster, Pennsylvania 17604, USA

Cancer radiotherapy uses high doses of ionizing radiation ($1-10^2$ Gy; 10^2-10^4 rad) because only a small fraction of the absorbed dose leads to lethal double-strand breaks in DNA. These breaks are more efficiently produced by Auger electrons ($1-10$ eV nm^{-1}) generated in proximity to the DNA. The energy of these electrons (on average 21 electrons for the decay of ^{125}I) is dissipated within 10-100 nm of the Auger event and produces multiple double-strand DNA breaks^{1,2}. A single Auger event can be lethal to a cell and is comparable to more than 10^5 photon absorption events in conventional radiotherapy^{3,4}. We now report that $^{57}\text{Fe(III)} \cdot \text{bleomycin}$, administered to malignant cells *in vitro* and *in vivo* and irradiated with resonant Mössbauer gamma rays (14.4 keV), causes ablation of the malignant cells, presumably by Auger cascade, with extremely small radiation doses—about 10^{-5} Gy. As a basis for comparison, about 5 Gy is necessary to achieve a similar effect with conventional radiotherapy⁵.

Bleomycin is one of many flat ring-shaped molecules that bind strongly to DNA by intercalating (inserting) between base pairs. The biologically active form of the drug is $\text{Fe(III)} \cdot \text{bleomycin}$. In this therapy (see the legend to Table 1 for the overall scheme) the photon-absorbing atom must be a Mössbauer isotope. Thus, we have prepared and used $^{57}\text{Fe(III)} \cdot \text{bleomycin}$ as our Mössbauer-isotope pharmaceutical.

After addition of the $^{57}\text{Fe(III)} \cdot \text{bleomycin}$ to the target cells, the cells were irradiated with highly monochromatic 14.4 keV gamma rays from excited-state ^{57}Fe nuclei (a Mössbauer isotope). In Mössbauer absorption, the radiation source is a collection of excited nuclei bound to a crystal lattice which emit highly monochromatic gamma rays as they decay to the ground state. The sample (absorber) consists of ground-state nuclei of the same isotope bound to a crystal lattice, which may or may not be the same lattice as the source lattice. In this case, ^{57}Co nuclei decay by electron capture to excited ^{57}Fe nuclei (the source nuclei), which emit highly monochromatic gamma-rays as they decay to the ground state: 137 keV, 123 keV and 14.4 keV. The absorbing nuclei are stable ground-state ^{57}Fe nuclei. When ^{57}Fe in the source and absorbing matrices is tightly bound, and when the source and absorbing matrices are the same, resonant absorption occurs (14.4 keV) because there is no recoil energy loss or gain by ^{57}Fe in either medium, and because the chemical environments are the same. However, if the energy states of the nuclei in the absorbing crystal are even slightly altered (by a different chemical environment or by an applied field, for example), the gamma rays will not be absorbed unless a Doppler velocity is applied to the source to compensate for the slightly

Table 1 The effect of treatment with drug $^{57}\text{Fe(III)}\cdot\text{bleomycin}$ and Mössbauer radiation on cell line HTB26 (human breast cancer tissue)

Experiment	Proliferation time (h)	Cell proliferation relative to no-treatment cells		
		No treatment	Drug alone	MIRAGE*
1	192	100	79	32
2	193	100	71	24
3	168	100	82	0
4	168	100	67	17
5	144	100	88	25
6	120	100	92	33
7	120	100	100	20

$^{57}\text{Fe(III)}\cdot\text{bleomycin}$ was prepared by dissolving ^{57}Fe metal (NEN) in concentrated HCl and neutralizing the solution with sodium hydroxide until Fe(OH)_3 just began to precipitate. Bleomycin (Bristol Myers) was added to the mixture (two Fe atoms per bleomycin). The pH of the yellow solution was adjusted to 7.4. The cells were grown in medium (Dulbecco's modified Eagles medium with 10% fetal bovine serum, $50\text{ }\mu\text{g ml}^{-1}$ streptomycin, $100\text{ }\mu\text{g ml}^{-1}$ vancomycin and 2 mM glutamine) in T25 flasks until a monolayer was obtained. Cells were treated in groups: (1) No-treatment cells—the monolayer of each flask was washed twice with iron-free medium (special order from Gibco); no drug or radiation was administered. (2) Drug alone—the monolayer of each flask was washed twice with iron-free medium; $50\text{ }\mu\text{l}$ drug $^{57}\text{Fe(III)}\cdot\text{bleomycin}$ ($2.3\times 10^{-4}\text{ M}$) was administered without radiation; the final concentration of $^{57}\text{Fe(III)}\cdot\text{bleomycin}$ was $3.8\text{ }\mu\text{M}$. (3) MIRAGE—the monolayer of each flask was washed twice with iron-free medium; $50\text{ }\mu\text{l}$ drug $^{57}\text{Fe(III)}\cdot\text{bleomycin}$ was administered and cells were irradiated for 65 min (5 h in experiment 3). After the time of the experiment had elapsed, the drug was removed by washing the monolayer twice with iron-free growth medium and once with phosphate-buffered saline. The cells were trypsinized with 5% trypsin EDTA and 10^5 cells from each experiment were passed into a new T25 flask containing a growth medium (counting was performed using methylene blue stain and a haemocytometer). The cells were grown as a monolayer for a period of time (120–192 h), after which they were trypsinized and counted a second time. The source (^{57}Co in a rhodium matrix) was driven at $+1.5\text{ mm s}^{-1}$, which is resonant with the Mössbauer absorption peak of this pharmaceutical¹⁸, by an Austin Science S-700 constant-velocity drive module in combination with an Austin Science K4 linear motor. The radiation dose was $8.3\times 10^{-5}\text{ Gy h}^{-1}$ (total) and $0.93\times 10^{-5}\text{ Gy h}^{-1}$ for the 14.4 keV gamma ray. The overall scheme is as follows. (1) Bleomycin is derivatized with $^{57}\text{Fe}^{3+}$ to give the drug, $^{57}\text{Fe(III)}\cdot\text{bleomycin}$. (2) The drug is administered to the target cells. The drug binds strongly to DNA. One drug molecule intercalates about every 8–10 base pairs. (3) The target cells are irradiated with 14.4 keV photons (about 10^{-5} Gy). (4) When an ^{57}Fe atom absorbs a 14.4 keV photon, the nucleus is transformed into an unstable excited state. About 90% of excited ^{57}Fe atoms decay by internal conversion (a nuclear rearrangement), followed by an Auger cascade (emission of many electrons). A molecular 'explosion' occurs in the vicinity because firstly the energy of the Auger electrons is dissipated within 10–100 nm, and secondly the cluster of positively charged atoms is blown apart by Coulombic repulsion as the highly charged Fe atom ionizes nearby atoms. (5) The Auger cascade and molecular explosion produce multiple double-strand breaks in the DNA.

* Velocity, 1.5 mm s^{-1} .

different nuclear energy levels in the two media. Extremely sharp lines that correspond to the natural lifetime of the excited state can be observed.

Of importance here is the fate of the excited ^{57}Fe absorber nuclei, which are tightly bound near the DNA. De-excitation occurs predominantly by internal conversion (a nuclear rearrangement) followed by an Auger cascade. In addition to the damage caused by the Auger electrons, the positive charge on the cascading ^{57}Fe atom (Fe^{+11} for eight Auger electrons) ionizes nearby atoms and causes a 'molecular explosion' because of the coulombic repulsion among this cluster of positively charged

atoms. It is not surprising that multiple double-strand breaks result when these events occur near DNA.

Resonant photon absorption by a Mössbauer isotope is much more probable than resonant photon absorption by other atoms in tissue, because the resonant-absorption cross-section for Mössbauer absorption (a nuclear excitation) is typically eight orders of magnitude greater than the total absorption cross-section of tissue. (The resonant-absorption cross-section, $2.2\times 10^{-17}\text{ cm}^2$ for Fe, should not be confused with the photoelectric cross-section, $5.5\times 10^{-21}\text{ cm}^2$ for Fe, which is typically three orders of magnitude less and requires radiation of a different multipolarity⁶.) Assuming one intercalated molecule (and one ^{57}Fe) per eight base pairs^{7–9}, 10^{10} base pairs per cell genome, and that one Mössbauer absorption per cell will be a lethal event, we calculate a required photon fluence of 3.8×10^7 photons cm^{-2} . Assuming that 60% of the photons are absorbed in 1 cm, we calculate a required dose (14.4 keV) of $5\times 10^{-5}\text{ Gy}$.

The effects of $1\times 10^{-5}\text{ Gy}$ doses of Mössbauer radiation on a human breast cancer cell line, HTB26, are shown in Table 1. The proliferation per cent in cell numbers was calculated as (final cell count of experiment)/(final cell count of No-treatment) $\times 100\%$. The ratio of the means of proliferation of the 'drug-alone' group (no Mössbauer radiation) and the MIRAGE (Mössbauer isotopic resonant absorption of gamma emission) group ($^{57}\text{Fe(III)}\cdot\text{bleomycin}$ with Mössbauer radiation) was 4.0; that is, the MIRAGE therapy reduced the cell proliferation by a factor of 4. The difference of the means of cell proliferation of the drug-alone group and the MIRAGE group was 62% ($P<0.005$). Similar results were obtained for human lung cancer cell line A549 and human breast cancer cell line MCF7.

The results of the treatment of tumour-bearing C3H mice with drug alone ($^{57}\text{Fe(III)}\cdot\text{bleomycin}$, no Mössbauer radiation) and with MIRAGE therapy (drug in combination with Mössbauer radiation) are given in Tables 2 and 3 respectively. The inhibition of growth index was calculated as 87% using the formula

$$\left(1 - \frac{T_n/T_0}{C_n/C_0}\right) \times 100\%$$

Where T_n is the mean tumour volume of the MIRAGE group on day 17, T_0 is the initial mean tumour volume of the MIRAGE group, C_n is the mean tumour volume of the drug-alone group on day 17, and C_0 is the initial mean tumour volume of the drug-alone group. The mean final:initial tumour volume ratio of the drug-alone group is 14. The mean final:initial tumour volume ratio of the MIRAGE group is 1.9. Thus, the average tumour was 7.4 times larger in the drug-alone group than in the MIRAGE group (14/1.9). The difference of the means of the ratios of the final:initial tumour volume of each mouse of the MIRAGE group and the drug-alone group is 12.1 ($P<0.005$). The MIRAGE group showed no weight change, whereas the drug-alone group demonstrated a significant increase in weight attributable to the increase in tumour bulk.

We have demonstrated a statistically significant reduction in cell proliferation in cell lines HTB26, A549 and MCF7, and in tumour bulk of C3H mice treated with intercalated $^{57}\text{Fe(III)}\cdot\text{bleomycin}$ and about 10^{-5} Gy of 14.4 keV radiation, which is of no consequence biologically. It is important to note that the Mössbauer isotope-carrying compound need not be toxic as a result of chemical or biological reactivity. Although bleomycin is toxic^{10,11}, it was used here (below the toxicity level) because all the appropriate parameters such as toxicity, binding constants, and Doppler shift for the Mössbauer effect are known. The ideal drug would have no toxicity at all; it would simply bind the Mössbauer isotope to the DNA. The drug concentration that achieves a level of intercalation sufficient for this treatment can be three orders of magnitude lower than the LD_{50} in mice^{12–17}. Furthermore, a conventional chemotherapeutic agent relies on a specific mechanism that constrains the effective structure of the agent and makes it subject to acquired tumour

Table 2 Effect of drug $^{57}\text{Fe(III)}$ ·bleomycin (no radiation) on C3H mice bearing spontaneous mammary adenocarcinomas

Mouse number	Initial tumour length (mm)	Initial tumour width (mm)	Final tumour length (mm)	Final tumour width (mm)	Initial tumour volume* (mm ³)	Final tumour volume* (mm ³)	Final vol/initial vol	Initial weight (g)	Weight change (g)
15	10.0	8.0	18.3	14.6	320	1,950	6.09	29.4	1.4
16	8.0	5.4	22.3	17.5	117	3,415	29.2	29.1	2.9
17	5.0	5.0	14.5	11.1	62	893	14.3	27.0	-0.5
18	8.3	8.0	22.0	20.7	266	4,713	17.7	32.0	0.6
19	7.0	6.1	23.4	22.5	130	5,936	45.7	31.1	2.6
20	9.4	8.7	17.1	15.4	356	2,028	5.70	28.8	0.5
21	10.2	10.1	17.3	15.1	520	1,972	3.79	29.4	4.6
22	10.4	7.0	20.0	17.1	255	2,924	11.5	29.7	4.4
23	9.5	7.0	17.4	12.4	236	1,349	5.72	29.6	0
24	9.8	6.7	19.8	17.5	220	3,024	13.7	25.5	4.5
25	7.8	4.9	17.0	16.7	94	2,371	25.2	27.5	0.9
26	6.2	5.4	9.2	6.4	90	190	2.12	31.5	-0.3
27	7.7	5.5	13.3	8.5	116	480	4.14	32.5	0.3
Average 14.4 ± 12 (s.d.)									1.7

All mice were anaesthetized with an intramuscular injection of 40 mg kg⁻¹ ketamine, 5 mg kg⁻¹ xylazine, and 0.03 mg kg⁻¹ atropine. Four mg kg⁻¹ drug $^{57}\text{Fe(III)}$ ·bleomycin was injected intratumorally (about 0.1 ml of 1.5 mg drug per ml). The mice were injected on days 1, 5, 9 and 13 and killed on day 17.

*Tumour volume was calculated as length $\times$ width²/2.

Table 3 Effect of $^{57}\text{Fe(III)}$ ·bleomycin with Mössbauer radiation on C3H mice bearing spontaneous mammary adenocarcinomas

Mouse number	Initial tumour length (mm)	Initial tumour width (mm)	Final tumour length (mm)	Final tumour width (mm)	Initial tumour volume* (mm ³)	Final tumour volume* (mm ³)	Final vol/initial vol	Initial weight (g)	Weight change (g)
1	10.5	10.5	9.1	8.0	579	291	0.503	29.0	-1.1
2	10.1	10.1	1.0	1.0	515	0.5	0.001	25.5	1.4
3	8.0	8.0	4.5	4.3	256	41	0.163	29.5	1.0
4	8.1	8.0	6.0	5.9	259	104	0.402	26.5	2.7
5	10.0	5.6	7.2	7.2	157	187	1.19	30.1	-0.8
6†	6.8	6.8	16.1	13.2	157	1,403	8.9	28.5	1.57
7	5.0	5.0	6.0	5.7	62	97	1.56	30.1	-3.5
8	13.0	11.2	11.8	9.9	811	578	0.713	31.5	-1.5
9	11.4	7.3	15.0	11.9	304	1,062	3.49	28.3	-0.5
10	10.6	7.2	11.8	10.0	277	593	2.14	30.8	-0.9
11	8.5	7.3	6.9	6.1	226	128	0.566	28.1	0.7
12	9.1	7.2	11.8	10.6	236	663	2.81	29.9	0.1
13	8.0	7.6	9.6	8.8	243	374	1.54	28.3	0.9
14	6.3	5.1	8.6	6.7	82	193	2.35	26.5	-1.0
Average 1.9 ± 2.2 (s.d.)									-0.07

All mice were anaesthetized with an intramuscular injection of 40 mg kg⁻¹ ketamine, 5 mg kg⁻¹ xylazine, and 0.03 mg kg⁻¹ atropine. Four mg kg⁻¹ drug $^{57}\text{Fe(III)}$ ·bleomycin was injected intratumorally (about 0.1 ml of 1.5 mg drug per ml). The tumour was placed directly under the source and irradiated for 15 min. Mice were treated on days 1, 5, 9 and 13. The radiation dose was $\sim 1 \times 10^{-4}$ Gy per treatment (total) and $\sim 1 \times 10^{-5}$ Gy per treatment for the 14.4 keV gamma ray. The mice were killed on day 17.

* Tumour volume was calculated as in Table 2.

† Tumour was difficult to immobilize because of its location on the chest wall.

resistance. In this therapy, tumour resistance can be ameliorated because any Mössbauer-isotope molecule that binds tightly to the DNA should be effective.

Negligible radiation doses and pharmacological nontoxicity together have important implications in human therapy for the elimination of pathological cell populations. Furthermore, the ability to control the occurrence of the Mössbauer effect over small spatial dimensions by manipulation of the resonance conditions, with the application of a magnetic field for example, could be a basis for selective cell-eradication therapy in humans.

Received 24 June; accepted 10 November 1988.

- Charlton, D. E. & Booz, J. *Radiat. Res.* **87**, 10-23 (1981).
- Linz, U. & G. Stoecklin, G. *7th International Congress of Radiation Research* (Amsterdam, 1983).
- Hofer, K. G. & Hughes, W. L. *Radiat. Res.* **47**, 94-109 (1971).
- Feinendegen, L. E., Ertl, H. H. & Bond, V. P. *Proc. Symp. Biophysical Aspects of Radiation Quality* 419-430 (International Atomic Energy Agency, Australia 1971).
- Hall, E. J. *Radiology for the Radiologist* 2nd edn, 225 (Harper & Row, Philadelphia, 1978).
- Hannon, J. P., Carron, N. J. & Trammell, G. T. *Phys. Rev.* **B9**, 2791-2809 (1974).
- Paoletti, J., Magee, B. B. & Magee, P. T. *Biochemistry* **16**, 351-357 (1977).
- Sakai, T. T., Riordan, J. M., Booth, T. E. & Glickson, J. D. *J. Med. Chem.* **24**, 279-285 (1981).
- Povirk, L. F., Hogan, M. & Dattagupta, N. *Biochemistry* **18**, 96-101 (1979).
- Rao, E. A., Saryan, L. A., Antholine, W. E. & Petering, D. H. *J. Med. Chem.* **23**, 1310-1318 (1980).
- Lin, P., Kwock, L., Hefter, K. & Misslbeck, G. *Cancer Res.* **43**, 1049-1053 (1983).
- Wu, H., N. Dattagupta, M. Hogan & Crothers, D. M. *Biochemistry* **19**, 626-634 (1980).
- Lawrence, J. & Daune, M. *Biochemistry* **15**, 3301-3307 (1976).
- Lawrence, J., Chan, D. C. F. & Piette, L. H. *Nucleic Acids Res.* **3**, 2879-2893 (1976).
- Paoletti, C. *et al. Recent Results in Cancer Research* **74**, 107-122 (1980).
- Stecker, P. G. (ed) *The Merck Index* 8th edn, 6 (Merck & Co., New Jersey, 1968).
- Clarke, E. G. C. *Isolation and Identification of Drugs* Vol. 1, 366 (Pharmaceutical Press, London, 1969).
- Burger, R. M., Kent, T. A., Horwitz, S. B., Munck, E. & Peisach, J. *J. Biol. Chem.* **258**, 1559-1564 (1983).

Correlation of DNA adduct levels in human lung with cigarette smoking

David H. Phillips*, Alan Hewer*, Carl N. Martin†, R. Colin Garner† & Maureen M. King†

* Chester Beatty Laboratories, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

† Cancer Research Unit, University of York, Heslington, York YO1 5DD, UK

Lung cancer is the most common cancer in men in the United Kingdom and the second most common in women, accounting for between 25 and 40% of all cancer deaths^{1,2}. Cigarette smoking is widely accepted as the major cause of lung cancer and linear relationships have been established between the number of cigarettes smoked and lung cancer risk³. Although approximately 50 carcinogenic chemicals have been identified in cigarette smoke⁴, a causal link between specific compounds and lung cancer has yet to be made. Studies on cigarette smokers' urine^{5,6}, blood⁷⁻¹¹ and placenta¹² have provided indications of carcinogen exposure, and although the presence of covalently-bound adducts in human DNA provides evidence of exposure to carcinogens¹²⁻¹⁴, there have been no reports of systematic studies on the levels of DNA adducts in human lung. We report here, using the ³²P-post-labelling technique^{15,16}, that cigarette smokers have higher adduct levels than non-smokers, that there is a linear relationship between adduct levels and daily or lifetime cigarette consumption, and that people who have given up smoking for at least five years have adduct levels similar to those of non-smokers.

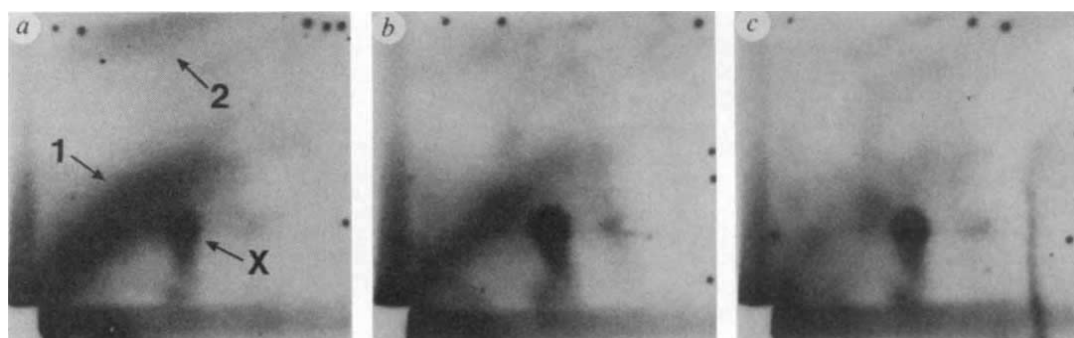
Coded lung samples were obtained from patients undergoing thoracic surgery at Bradford Royal Infirmary. Patients were asked whether or not they smoked cigarettes, how many they smoked each day and for how many years. Where a patient did not indicate how long he or she had smoked, we assumed it was since the age of 18. DNA isolated from non-tumorous lung tissue was hydrolysed to 3' mononucleotides, 5'-labelled with high specific activity [γ -³²P]ATP, and analysed by thin-layer chromatography. The detection and measurement of DNA adducts using the ³²P-post-labelling technique relies on differences in chromatographic mobility between normal and carcinogen-modified nucleotides, and levels as low as 1 in 10⁹-10¹⁰ nucleotides can be detected for aromatic and other hydrophobic adducts.

Autoradiographs of chromatograms of ³²P-post-labelled digests of DNA from smokers' lungs revealed a strong band of radioactive material (Fig. 1a arrow '1') and a weaker, faster migrating band (Fig. 1a '2'). This pattern is similar to that found in lung DNA from mice treated topically with cigarette smoke condensate¹⁷. A radioactive spot (Fig. 1a 'X') close to the stronger band is a background spot that is observed in all DNA samples, regardless of their tissue or species origin^{17,18}. The adduct pattern observed for an ex-smoker of more than five years abstinence (Fig. 1b) was qualitatively similar but the radioactive bands were much weaker. A former smoker who had given up three months before surgery had a pattern of adduct intensities similar to those of the smokers (data not shown). The typical adduct pattern of a non-smoker contained no strong adduct bands (Fig. 1c), although low levels of radioactive material were detected in the same region of the chromatogram as band 1 in the smokers' samples.

Of 29 samples analysed, five were from non-smokers and seven were from former smokers. Table 1 give the individual patients' details and smoking histories. Kreyberg Group I tumours, that is, those more commonly related to smoking¹⁹, accounted for 11/17 of the smokers', 6/7 of the former smokers' and 2/5 of the non-smokers' tumours. Total adduct levels in the smokers' DNA samples, estimated from the levels of radioactivity in the adduct bands in the chromatograms, ranged from 1.5 to 34.3 adducts per 10⁸ nucleotides, with 12/17 samples having levels in excess of five adducts per 10⁸ nucleotides. The weaker band 2 (Fig. 1a) accounted for less than 5% of the total adducts on the chromatograms. The samples from non-smokers had 1.9-3.3 adducts per 10⁸ nucleotides. Of the seven former smokers, those who had given up smoking only 1-3 months previously (YT8 and YT25) had adduct levels typical of the smokers (13.5 and 12.5 adducts per 10⁸ nucleotides, respectively) whereas 3/4 who gave up more than five years ago had adduct levels similar to the non-smokers (1.7-4.9 adducts per 10⁸ nucleotides). Regression analysis of the data from the smokers revealed a linear relationship between the number of cigarettes smoked per day and the DNA adduct levels (correlation coefficient = 0.724, $P < 0.001$, 20 degrees of freedom) (Fig. 2). Plotting adduct levels against estimated lifetime cigarette consumption also showed a linear relationship (data not shown), but the correlation is not as strong (correlation coefficient = 0.663, $P < 0.001$, 20 degrees of freedom).

These results show, for the first time, a linear relationship between numbers of cigarettes smoked and DNA adduct levels in the lung, as measured by the ³²P-post-labelling method. In

Fig. 1 Autoradiographs of chromatograms of ³²P-post-labelled digests of human lung DNA isolated from a, a cigarette smoker (20 cigarettes per day, sample YT27); b, a former smoker who stopped smoking five years previously (sample YT29); and c, a non-smoker (sample YT30). The origins, located at the bottom left-hand corner of the chromatograms, were excised before autoradiography.



Methods. DNA was isolated from 1-5 g lung tissue and purified as described elsewhere²⁰. DNA samples (4 μ g) were digested with micrococcal nuclease and spleen phosphodiesterase, and then with nuclease P₁ as described¹⁸. The DNA digest was then labelled with [γ -³²P]ATP (75 μ Ci, 1000 Ci mmol⁻¹) and the reaction terminated with apyrase¹⁵. Resolution of the ³²P-labelled adducts was carried out on polyethyleneimine-cellulose TLC sheets (10 $\times$ 10 cm) which were developed overnight with 1 M sodium phosphate, pH 6.0, on to a filter-paper wick¹² and then, after removal of the wick, in the reverse direction with 3.5 M lithium formate, 8.5 M urea, pH 3.5. The sheets were turned through 90° and developed first with 0.8 M LiCl, 0.5 M tris-HCl, 8.5 M urea, pH 8.0 and then with 1.7 M sodium phosphate, pH 6.0, in the latter case using a filter-paper wick. The chromatograms were then visualized by autoradiography for three days at -70 °C using intensifying screens. Radioactive ink (seen as small spots at the edges of the chromatograms) was used to align the developed autoradiographs with the TLC sheets. The levels of adducts were determined by excising areas of the chromatograms and determining the levels of radioactivity present by Cerenkov counting¹⁶.

Table 1 DNA adducts in human lung

Patient	Age	Sex	Smoking history	Tumour	Kreyberg group	Cigarettes per day	Total packets ^h ($\times 10^4$)	Adducts per 10^8 nucleotides ⁱ
YT1	70	M	FS ^a	A	II	—	1.4	4.9 $\pm$ 2.6
YT2	66	F	S	uncharacterized		40	3.5	18.3 $\pm$ 6.5
YT3	71	M	S	SQ	I	10	0.9	2.4 $\pm$ 0.0
YT4	72	F	S	SQ	I	20	1.97	3.1 $\pm$ 0.9
YT5	51	M	S	SQ	I	25	1.20	8.9 $\pm$ 1.6
YT6	58	M	S	SQ	I	30	2.19	18.4 $\pm$ 2.2
YT7	71	M	NS	BA	II	—	—	2.0 $\pm$ 0.3
YT8	54	M	FS ^b	SQ	I	25	1.3	13.5 $\pm$ 3.9
YT9	62	M	S	SQ	I	20	1.6	1.5 $\pm$ 0.1
YT10	53	M	FS ^c	SQ	I	—	0.9	12.1 $\pm$ 1.0
YT11	55	F	S	PA	II	15-20	1.3	10.7 $\pm$ 1.0
YT12	77	M	S	BA	II	10	1.0	3.1 $\pm$ 0.0
YT13	60	M	NS	SQ	I	—	—	3.0 $\pm$ 1.3
YT14	69	F	S	SQ	I	10-15	1.93	2.7 $\pm$ 0.1
YT15	74	M	FS ^d	SQ	I	—	1.6	1.7 $\pm$ 0.5
YT16	63	F	NS	A	II	—	—	2.1 $\pm$ 0.3
YT17	62	M	S	SQ	I	15-20	1.8	16.6 $\pm$ 0.2
YT18	66	F	S	SQ	I	20	1.46	13.1 $\pm$ 1.3
YT19	68	M	S	A	II	10	0.7	16.4 $\pm$ 4.1
YT20	61	M	S	SQ	I	20	1.5	17.6 $\pm$ 1.9
YT21	56	M	FS ^e	SQ	I	20	1.3	15.2 $\pm$ 0.9
YT22	64	M	S	LC	I/II	40	3.35	34.3 $\pm$ 4.0
YT23	61	M	S	C	II	5-6	0.2	7.7 $\pm$ 0.8
YT24	66	F	S	SQ	I	30	2.6	9.3 $\pm$ 1.9
YT25	55	M	FS ^f	CS	I	20	1.4	12.5 $\pm$ 2.6
YT26	61	M	NS	SQ	I	—	—	3.3 $\pm$ 0.0
YT27	52	F	S	SQ	I	20	1.24	8.4 $\pm$ 0.3
YT29	65	M	FS ^g	SQ	I	—	1.71	2.3 $\pm$ 0.8
YT30	48	F	NS	LC	I/II	—	—	1.9 $\pm$ 0.5

FS, former smoker, NS, non-smoker; S, smoker. A, adenocarcinoma; SQ, squamous-cell carcinoma; BA, bronchiolo-alveolar carcinoma; PA, papillary adenocarcinoma; LC, poorly differentiated large-cell carcinoma, mixed squamous/adenocarcinoma; C, carcinoid tumour; CS, carcinosarcoma. *a*, Stopped smoking 1976. *b*, Stopped smoking 1 month before operation. *c*, Stopped smoking 1977. *d*, Stopped smoking 1976. *e*, Stopped smoking, date unknown (recent?). *f*, Stopped smoking 3 months before operation. *g*, Stopped smoking 1982. *h*, Total packets smoked before operation. One packet is 20 cigarettes. *i*, Mean $\pm$ range of 2-3 determinations performed on separate days.

addition, they demonstrate the loss of adducts after smoking ceased but show that this process takes more than three months. The chromatographic profiles observed are broadly characteristic of adducts formed by polycyclic aromatic compounds¹⁵. The use of nuclease P₁ digestion to enhance analytical sensitivity could result, however, in the dephosphorylation of certain adducts, such as those formed by aromatic amines, leading to their failure to be ³²P-labelled and to an underestimation of the true values of the adduct levels. Smaller or more polar adducts, for example those formed by tobacco-specific nitrosamines²⁰, might also be dephosphorylated or lost during the development of the chromatograms and thus go undetected. Despite these

uncertainties in the absolute quantitation of adduct levels, relative values of the determinations are not in doubt. Clear resolution of the adducts into discrete spots was not achieved, suggesting that a large number of different compounds were covalently bound to DNA. Similar adduct profiles have been obtained following the application of carcinogenic mixtures of polycyclic aromatic hydrocarbons to experimental animals^{17,21}.

No account has been taken of occupation, diet, socioeconomic group or environment in interpreting our findings. Nevertheless, the adduct pattern for all smokers was similar, differing only in intensity. The presence of detectable levels of adducts in non-smokers could be the result of passive smoking

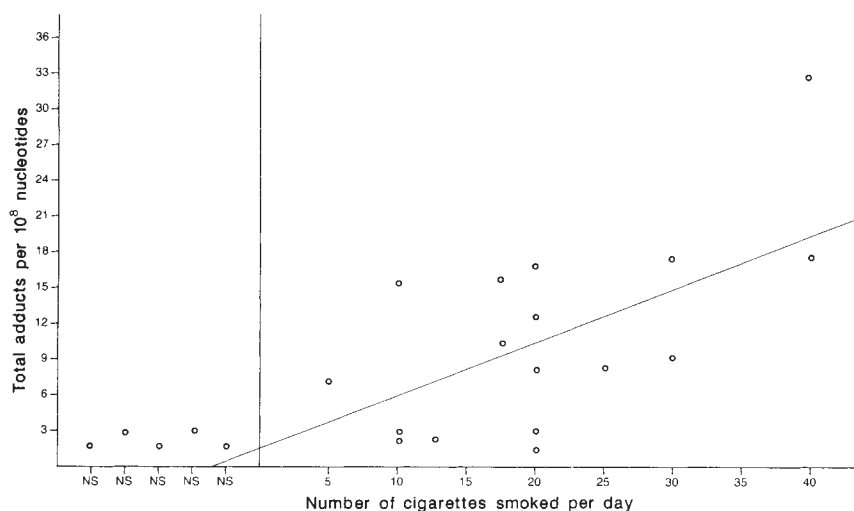


Fig. 2 Relationship between cigarettes smoked per day and lung DNA adduct levels estimated by ³²P-post-labelling. NS, non-smoker. Former smokers have been excluded from the analysis. Linear regression analysis gave the line of best fit shown, $y = 1.69 + 0.48x$ (correlation coefficient 0.724, $P < 0.001$, 20 degrees of freedom).

or of exposure to other DNA-damaging agents, not related to tobacco, that are present in the environment (for example fossil fuel combustion products), or both. Our results contrast with those of studies on peripheral blood lymphocytes, in which no difference in adduct levels between smokers and non-smokers was observed^{9,14,18,22}. This could be related to peripheral lymphocytes having efficient DNA repair pathways, or to the activated metabolites generated in lung cells not migrating from these cells and thus failing to react with lymphocyte DNA. Although it has been demonstrated that lymphocyte DNA from workers occupationally exposed to carcinogens shows the presence of adducts^{14,23,24}, for cigarette smoking, adduct measurement in lymphocyte DNA may not be a good indicator of what is happening in the bronchial epithelium, the site of smoking-induced cancer.

The levels of DNA adducts in the lungs of a cigarette smoker are probably the result of a balance between the formation of new adducts and their loss through cell turnover and DNA repair processes. Thus a steady state may be achieved which would be expected to be subject to some variation between individuals^{25,26}. The demonstration of a linear relationship between daily cigarette consumption and adduct levels lends powerful support to the epidemiological finding that the more cigarettes smoked the greater the risk of lung cancer. Although it is not known which, if any, of the adducts measured is causally involved in the process of lung carcinogenesis, it is clear that DNA binding of the putative carcinogens is, as has been demonstrated in experimental animal studies²⁷, linearly related to dose. Of comfort to ex-smokers is the fact that stopping smoking can lead to a loss of DNA adducts, even though this process may take months to years. Although the reduced risk of lung cancer after smoking has ended²⁸ has been interpreted as indicating an interruption in a late stage of the carcinogenic process, it may be also due to the loss of the pro-mutagenic lesions that initiate the process.

We thank Mr A. J. Mearns, Consultant Thoracic Surgeon, Bradford Royal Infirmary, for providing lung tissue samples. This study was supported in part by a research contract from the UK Health and Safety Executive and, in part, by the MRC and the CRC.

Received 13 June; accepted 11 November 1988.

1. Doll, R. & Hill, A. B. *Br. med. J.* **ii**, 1071-1081 (1956).
2. *Yorkshire Regional Cancer Registry: Report for the Year 1985 Including Cancer Statistics for 1978-1983* (Yorkshire Regional Cancer Organisation, Leeds, 1986).
3. Doll, R. & Hill, A. B. *Br. med. J.* **ii**, 1271-1286 (1952).
4. Weisburger, J. H., Cohen, L. A. & Wynder, E. L. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winsten, J. A.) 567-602 (Cold Spring Harbor, New York, 1977).
5. Putzrath, R. M., Langden, D. & Eisenstadt, E. *Mutat. Res.* **85**, 97-108 (1981).
6. Yamasaki, E. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3555-3559 (1977).
7. Obe, G., Vogt, H.-J., Madle, S., Fahning, A. & Heller, W. D. *Mutat. Res.* **92**, 303-319 (1982).
8. Livingston, G. K. & Fineman, R. M. *Mutat. Res.* **119**, 59-64 (1983).
9. Perera, F. P. *et al. J. natn. Cancer Inst.* **79**, 449-456 (1987).
10. Tornqvist, M. *et al. Carcinogenesis (Lond.)* **7**, 1519-1521 (1986).
11. Bryant, M. S., Skipper, P. L., Tannenbaum, S. R. & Maclure, M. *Cancer Res.* **47**, 602-608 (1987).
12. Everson, R. B. *et al. Science* **231**, 54-57 (1986).
13. Garner, R. C. *Carcinogenesis (Lond.)* **6**, 1071-1078 (1985).
14. Phillips, D. H., Hemminki, H., Alhonen, A., Hewer, A. & Grover, P. L. *Mutat. Res.* **204**, 531-541 (1988).
15. Gupta, R. C., Reddy, M. V. & Randerath, K. *Carcinogenesis (Lond.)* **3**, 1081-1092 (1982).
16. Reddy, M. V. & Randerath, K. *Carcinogenesis (Lond.)* **7**, 1543-1551 (1986).
17. Randerath, E., Mittal, D. & Randerath, K. *Carcinogenesis (Lond.)* **9**, 75-80 (1988).
18. Phillips, D. H., Hewer, A. & Grover, P. L. *Carcinogenesis (Lond.)* **7**, 2071-2075 (1986).
19. *International Agency for Research on Cancer Monographs on the Carcinogenic Risk of Chemical to Humans Vol. 38, Tobacco Smoking* (IARC, Lyon 1986).
20. Hecht, S. S. & Hoffmann, D. *Carcinogenesis (Lond.)* **9**, 875-884 (1988).
21. Schoket, B., Hewer, A., Grover, P. L. & Phillips, D. H. *Carcinogenesis (Lond.)* **9**, 1253-1258 (1988).
22. Everson, R. B. *et al. Prog. exp. Tumor Res.* **31**, 86-103 (1987).
23. Harris, C. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 6672-6676 (1985).
24. Perera, F. P. *et al. Cancer Res.* **48**, 2288-2291 (1988).
25. Harris, C. C. *et al. Cancer Res.* **36**, 1011-1018 (1976).
26. Sabadie, N., Richter-Reichhelm, H. B., Saracci, R., Mohr, U. & Bartsch, H. *Int. J. Cancer* **27**, 417-425 (1979).
27. Pereira, M. A., Burns, F. J. & Albert, R. E. *Cancer Res.* **39**, 2556-2559 (1979).
28. Doll, R. *Cancer Res.* **38**, 3573-3583 (1978).
29. Gupta, R. C. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6943-6947 (1984).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article
Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
 Title _____
 Institution/Company _____
 Department _____
 Address _____
 City _____ State _____ Zip _____
 Phone (_____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106